

IJCSIS Vol. 8 No. 7, October 2010
ISSN 1947-5500

International Journal of Computer Science & Information Security

© IJCSIS PUBLICATION 2010

Editorial Message from Managing Editor

IJCSIS is a monthly and open access publishing venue for research in general computer science and information security. This journal, Vol. 8 No. 7 October 2010 issue, is popular and highly esteemed among research academics, university IT faculties; industry IT departments; government departments; the mobile industry and computing industry.

The aim is to publish high quality papers on a broad range of topics: security infrastructures, network security: Internet security, content protection, cryptography, all aspects of information security; computer science, computer applications, multimedia systems, software engineering, information systems, intelligent systems, web services, data mining, wireless communication, networking and technologies, innovative technology and management.

Original research and state-of-the-art results has in this journal a natural home and an outlet to the broader scientific community. I congratulate all parties responsible for supporting this journal, and wish the editorial board & technical review committee a successful operation.

Available at <http://sites.google.com/site/ijcsis/>

IJCSIS Vol. 8, No. 7, October 2010 Edition

ISSN 1947-5500 © IJCSIS, USA.

Indexed by Google Scholar, ESBCOHOST, ProQuest, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Cornell University Library, ScientificCommons, and more.



IJCSIS EDITORIAL BOARD

Dr. Gregorio Martinez Perez

Associate Professor - Professor Titular de Universidad, University of Murcia (UMU), Spain

Dr. M. Emre Celebi,

Assistant Professor, Department of Computer Science, Louisiana State University in Shreveport, USA

Dr. Yong Li

School of Electronic and Information Engineering, Beijing Jiaotong University, P. R. China

Prof. Hamid Reza Naji

Department of Computer Engineering, Shahid Beheshti University, Tehran, Iran

Dr. Sanjay Jasola

Professor and Dean, School of Information and Communication Technology, Gautam Buddha University

Dr Riktesh Srivastava

Assistant Professor, Information Systems, Skyline University College, University City of Sharjah, Sharjah, PO 1797, UAE

Dr. Siddhivinayak Kulkarni

University of Ballarat, Ballarat, Victoria, Australia

Professor (Dr) Mokhtar Beldjehem

Sainte-Anne University, Halifax, NS, Canada

Dr. Alex Pappachen James, (Research Fellow)

Queensland Micro-nanotechnology center, Griffith University, Australia

Dr. T.C. Manjunath,

ATRIA Institute of Tech, India.

TABLE OF CONTENTS

1. Paper 29091048: Data Group Anonymity: General Approach (pp. 1-8)

*Oleg Chertov, Applied Mathematics Department, NTUU "Kyiv Polytechnic Institute", Kyiv, Ukraine
Dan Tavrov, Applied Mathematics Department, NTUU "Kyiv Polytechnic Institute", Kyiv, Ukraine*

2. Paper 26091026: A Role-Oriented Content-based Filtering Approach: Personalized Enterprise Architecture Management Perspective (pp. 9-18)

*Imran GHANI, Choon Yeul LEE, Seung Ryul JEONG, Sung Hyun JUHN
(School of Business IT, Kookmin University, Seoul 136-702, Korea)
Mohammad Shafie Bin Abd Latiff
(Faculty of Computer Science and Information Systems Universiti Teknologi Malaysia, 81310, Malaysia)*

3. Paper 28091036: Minimizing the number of retry attempts in keystroke dynamics through inclusion of error correcting schemes (pp. 19-25)

*Pavaday Narainsamy, Student member IEEE, Computer Science Department, Faculty of Engineering, University Of Mauritius
Professor K.M.S.Soyjaudah, Member IEEE, Faculty of Engineering, University of Mauritius*

4. Paper 29091049: Development of Cinema Ontology: A Conceptual and Context Approach (pp. 26-31)

*Dr. Sunitha Abburu, Professor, Department of Computer Applications, Adhiyamaan College of Engineering, Hosur, India
Jinesh V N, Lecturer, Department of Computer Science, The Oxford College of Science, Bangalore, India*

5. Paper 13091002: S-CAN: Spatial Content Addressable Network for Networked Virtual Environments (pp. 32-38)

*Amira Soliman, Walaa Sheta
Informatics Research Institute, Mubarak City for Scientific Research and Technology Applications, Alexandria, Egypt.*

6. Paper 26091024: Combinatory CPU Scheduling Algorithm (pp. 39-43)

*Saeeda Bibi, Farooque Azam,
Department of Computer Engineering, College of Electrical and Mechanical Engineering, National University of Science and Technology, Islamabad, Pakistan
Yasir Chaudhry, Department of Computer Science, Maharishi University of Management, Fairfield, Iowa USA*

7. Paper 26091046: Enterprise Crypto method for Enhanced Security over semantic web (pp. 44-48)

*Talal Talib Jameel, Department of Medical Laboratory Science s, Al Yarmouk University College
Baghdad, Iraq*

8. Paper 30091054: On the Performance of Symmetrical and Asymmetrical Encryption for Real-Time Video Conferencing System (pp. 49-55)

*Maryam Feily, Salah Noori, Sureswaran Ramadass
National Advanced IPv6 Centre of Excellence (NAv6), Universiti Sains Malaysia (USM), Penang, Malaysia*

9. Paper 11101004: RACHSU Algorithm based Handwritten Tamil Script Recognition (pp. 56-61)

*C. Sureshkumar, Department of Information Technology, J.K.K.Nataraja College of Engineering, Namakkal, Tamilnadu, India
Dr. T. Ravichandran, Department of Computer Science & Engineering, Hindustan Institute of Technology, Coimbatore, Tamilnadu, India*

10. Paper 13081003: Trust challenges and issues of E-Government: E-Tax prospective (pp. 62-66)

Dinara Berdykhanova, Asia Pacific University College of Technology and Innovation Technology Park Malaysia, Kuala Lumpur, Malaysia

Ali Dehghantanha, Asia Pacific University College of Technology and Innovation Technology Park Malaysia, Kuala Lumpur- Malaysia

Andy Seddon, Asia Pacific University College of Technology and Innovation Technology Park Malaysia, Kuala Lumpur- Malaysia

11. Paper 16081008: Machine Learning Approach for Object Detection - A Survey Approach (pp. 67-71)

N.V. Balaji, Department of Computer Science, Karpagam University, Coimbatore, India

Dr. M. Punithavalli, Department of Computer Science, Sri Ramakrishna Arts College for Women, Coimbatore, India

12. Paper 18061028: Performance comparison of SONET, OBS on the basis of Network Throughput and Protection in Metropolitan Networks (pp. 72-75)

Mr. Bhupesh Bhatia, Assistant Professor , Northern India Engineering College, New Delhi, India

R.K.Singh, Officer on special duty, Uttarakhand Technical University, Dehradun (Uttarakhand), India

13. Paper 23091017: A Survey on Session Hijacking (pp. 76-83)

P. Ramesh Babu, Dept of CSE, Sri Prakash College of Engineering, Tuni-533401, INDIA

D. Lalitha Bhaskari, Dept of CS & SE, AU College of Engineering (A), Visakhapatnam-530003, INDIA

CPVNJ Mohan Rao, Dept of CSE, Avanthi Institute of Engineering & Technology, Narsipatnam-531113, INDIA

14. Paper 26091022: Point-to-Point IM Interworking session Between SIP and MFTS (pp. 84-87)

Mohammed Faiz Aboalmaaly, Omar Amer Abouabdalla, Hala A. Albaroodi and Ahmed M. Manasrah

National Advanced IPv6 Centre, Universiti Sains Malaysia, Penang, Malaysia

15. Paper 29071042: An Extensive Survey on Gene Prediction Methodologies (pp. 88-104)

Manaswini Pradhan, Lecturer, P.G. Department of Information and Communication Technology, Fakir Mohan University, Orissa, India

Dr. Ranjit Kumar Sahu, Assistant Surgeon, Post Doctoral Department of Plastic and Reconstructive Surgery, S.C.B. Medical College, Cuttack, Orissa, India

16. Paper 29091040: A multicast Framework for the Multimedia Conferencing System (MCS) based on IPv6 Multicast Capability (pp. 105-110)

Hala A. Albaroodi, Omar Amer Abouabdalla, Mohammed Faiz Aboalmaaly and Ahmed M. Manasrah

National Advanced IPv6 Centre, Universiti Sains Malaysia, Penang, Malaysia

17. Paper 29091042: The Evolution Of Chip Multi-Processors And Its Role In High Performance And Parallel Computing (pp. 111-117)

A. Neela madheswari, Research Scholar, Anna University, Coimbatore, India

Dr. R.S.D. Wahida banu, Research Supervisor, Anna University, Coimbatore, India

18. Paper 29091044: Towards a More Mobile KMS (pp. 118-123)

Julius Olatunji Okesola, Dept. of Computer and Information Sciences, Tai Solarin University of Education, Ijebu-Ode, Nigeria

Oluwafemi Shawn Ogunseye, Dept. of Computer Science, University of Agriculture, Abeokuta, Nigeria

Kazeem Idowu Rufai, Dept. of Computer and Information Sciences, Tai Solarin University of Education, Ijebu-Ode, Nigeria

19. Paper 30091055: An Efficient Decision Algorithm For Vertical Handoff Across 4G Heterogeneous Wireless Networks (pp. 124-127)

S. Aghalya, P. Seethalakshmi,

Anna University Tiruchirappalli, India

20. Paper 231010XX: Combining Level- 1 ,2 & 3 Classifiers For Fingerprint Recognition System (pp. 128-132)

*Dr. R. Seshadri , B.Tech, M.E,Ph.D, Director, S.V.U.Computer Center, S.V.University, Tirupati
Yaswanth Kumar.Avulapati, M.C.A,M.Tech,(Ph.D), Research Scholar, Dept of Computer Science,
S.V.University, Tirupati*

21. Paper 251010XX: Preventing Attacks on Fingerprint Identification System by Using Level-3 Features (pp. 133-138)

*Dr. R. Seshadri , B.Tech, M.E,Ph.D, Director, S.V.U.Computer Center, S.V.University, Tirupati
Yaswanth Kumar.Avulapati, M.C.A,M.Tech,(Ph.D), Research Scholar, Dept of Computer Science,
S.V.University, Tirupati*

22. Paper 13091003: Using Fuzzy Support Vector Machine in Text Categorization Base on Reduced Matrices (pp. 139-143)

Vu Thanh Nguyen, University of Information Technology HoChiMinh City, VietNam

23. Paper 13091001: Categories Of Unstructured Data Processing And Their Enhancement (pp. 144-150)

*Prof.(Dr). Vinodani Katiyar, Sagar Institute of Technology and Management, Barabanki U.P. India.
Hemant Kumar Singh, Azad Institute of Engineering & Technology, Lucknow, U.P. India*

24. Paper 30091071: False Positive Reduction using IDS Alert Correlation Method based on the Apriori Algorithm (pp. 151-155)

*Homam El-Taj, Omar Abouabdalla, Ahmed Manasrah, Mohammed Anbar, Ahmed Al-Madi National
Advanced IPv6 Center of Excellence (NAv6) Universiti Sains Malaysia, Penang, Malaysia*

25. Paper 21091012: Sector Mean with Individual Cal and Sal Components in Walsh Transform Sectors as Feature Vectors for CBIR (pp. 156-164)

*Dr. H. B. Kekre, Senior Professor, Computer Engineering, MPSTME,SVKM'S NMIMS University, Mumbai,
India.
Dhirendra Mishra, Associate Professor, Computer Engineering, MPSTME, SVKM'S NMIMS University,
Mumbai, India.*

26. Paper 23091015: Supervised Learning approach for Predicting the Presence of Seizure in Human Brain (pp. 165-169)

*Sivagami P, Sujitha V, M.Phil Research Scholar, PSGR Krishnammal College for Women, Coimbatore,
India
Vijaya MS, Associate Professor and Head GRG School of Applied Computer Technology, PSGR
Krishnammal College for Women, Coimbatore, India.*

27. Paper 28091038: Approximate String Search for Bangla: Phonetic and Semantic Standpoint (pp. 170-174)

*Adeeb Ahmed, Department of Electrical and Electronic Engineering, Bangladesh University of
Engineering and Technology Dhaka, Bangladesh
Abdullah Al Helal, Department of Electrical and Electronic Engineering, Bangladesh University of
Engineering and Technology Dhaka, Bangladesh*

28. Paper 29091045: Multicast Routing and Wavelength Assignment for Capacity Improvement in Wavelength Division Multiplexing Networks (pp. 175-182)

*N. Kaliammal, Professor, Department of ECE, N.P.R college of Engineering and Technology, Dindugul,
Tamil nadu
G. Gurusamy, Dean/HOD EEE, FIE, Bannari amman Institute of Technology, Sathyamangalam,Tamil
nadu.*

29. Paper 30091056: Blind Robust Transparent DCT-Based Digital Image Watermarking for Copyright Protection (pp. 183-188)

Hanan Elazhary and Sawsan Morkos

Computers and Systems Department, Electronics Research Institute, Cairo, Egypt

30. Paper 25091019: An Enhanced LEACH Protocol using Fuzzy Logic for Wireless Sensor Networks (pp. 189-194)

J. Rathi, K. S. Rangasamy college of technology, Tiruchengode, Namakkal(Dt)-637 215, Tamilnadu, India

Dr. G. Rajendran, Kongu Engg. College, Perundurai, Erode(Dt)-638 052, Tamilnadu, India

31. Paper 29091050: A Novel Approach for Hiding Text Using Image Steganography (pp. 195-200)

Sukhpreet Kaur, Department of Computer Science and Engineering, Baba Farid College of Engineering and Technology, Bathinda-151001, Punjab, India

Sumeet Kaur, Department of Computer Engineering, Yadavindra College of Engineering Punjabi University Guru Kashi Campus, Talwandi Sabo, Punjab, India

32. Paper 30091058: An approach to a pseudo real-time image processing engine for hyperspectral imaging (pp. 201-207)

Sahar Sabbaghi Mahmoudi, Smart Technology and Robotics Programme, Institute of Advanced Technology (ITMA), Universiti Putra Malaysia, Serdang, Malaysia

Prof. Dr. Shattri Mansor, Remote Sensing and GIS Programme, Department of Civil Engineering, Universiti Putra Malaysia, Serdang, Malaysia

Abed Abedniya, MBA Programme, Faculty of Management (FOM), Multimedia University, Malaysia

33. Paper 23091016: Improved Computer Networks resilience Using Social Behavior (pp. 208-214)

Yehia H. Khalil^{1,2}, Walaa M. Sheta² and Adel S. Elmaghraby¹

¹Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY

²Informatics Research Institute, MUCST, Burg El Arab, Egypt

34. Paper 27091034: Mobile Embedded Real Time System (RTTCS) for Monitoring and Controlling in Telemedicine (pp. 215-223)

Dr. Dhuha Basheer Abdullah, Asst. Prof. / computer sciences Dept. College of Computers and Mathematics / Mosul University Mosul / Iraq

Dr. Muddather Abdul-Alaziz, Lecturer / Emergency Medicine Dept, Mosul College of Medicine, Mosul University Mosul / Iraq

Basim Mohammed, Asst. lecturer / computer center, Mosul University Mosul / Iraq

35. Paper 01111001: Automating the fault tolerance process in Grid Environment (pp. 224-230)

Inderpreet Chopra, Research Scholar, Thapar University Computer Science Department, Patiala, India

Maninder Singh, Associate Professor, Thapar University Computer Science Department, Patiala, India

36. Paper 01111002: A Computational Model for Bharata Natyam Choreography (pp. 231-233)

Sangeeta Jadhav, S.S Dempo College of Commerce and Economics, Panaji, Goa India.

Sasikumar, CDAC, Mumbai, India.

37. Paper 01111003: Haploid vs Diploid Genome in Genetic Algorithms for TSP (pp. 234-238)

Rakesh Kumar, Associate Professor, Department of Computer Science & Application, Kurukshetra University, Kurukshetra

Jyotishree, Assistant Professor, Department of Computer Science & Application, Guru Nanak Girls College, Yamuna Nagar

38. Paper 01111004: Context Based Personalized Search Engine For Online Learning (pp. 239-244)

Dr. Ritu Soni, Prof. & Head, DCSA, GNG College, Santpura, Haryana, India

Mrs. Preeti Bakshi, Lect. Computer Science, GNG College, Santpura, Haryana, India

39. Paper 01111005: Self-Healing In Wireless Routing Using Backbone Nodes (pp. 245-252)

Urvi Sagar 1 , Ashwani Kush 2

2 CSE, NIT KKR,

1 Comp Sci Dept, University College, Kurukshetra University India

40. Paper 01111006: Vectorization Algorithm for Line Drawing and Gap filling of Maps (pp. 253-258)

Ms. Neeti Daryal, Lecturer, Department of Computer Science, M L N College, Yamuna Nagar

Dr Vinod Kumar, Reader, Department of Mathematics, J.V.Jain College, Saharanpur

41. Paper 01111007: Simulation Modeling of Reactive Protocols for Adhoc Wireless Network (pp. 259-265)

Sunil Taneja, Department of Computer Science, Government Post Graduate College, Kalka, India

Ashwani Kush, Department of Computer Science, University College, Kurukshetra University, Kurukshetra, India

Amandeep Makkar, Department of Computer Science, Arya Girls College, Ambala Cantt, India

42. Paper 01111008: Media changing the Youth Culture: An Indian Perspective (pp. 266-271)

Prof. Dr. Ritu Soni, Head, Department of Computer Science, Guru Nanak Girls' College, Yamuna Nagar, Haryana, India-135003

Prof. Ms. Bharati Kamboj, Department of Physics, Guru Nanak Girls' College, Yamuna Nagar, Haryana, India-135003

43. Paper: Reliable and Energy Aware QoS Routing Protocol for Mobile Ad hoc Networks (pp. 272-278)

V.Thilagavathe, Lecturer, Department of Master of Computer Applications, Institute of Road & Transport Technology

K.Duraiswamy, Dean, K.S. Rangasamy College of Technology, Tiruchengode

44. Paper: A Dynamic Approach To Defend Against Anonymous DDoS Flooding Attacks (Pp. 279-284)

Mrs. R. Anurekha, Lecturer, Dept. of IT, Institute of Road and Transport Technology, Erode, Tamilnadu, India.

Dr. K. Duraiswamy, Dean, Department of CSE, K.S.Rangasamy College of Technology, Tiruchengode, Namakkal, Tamilnadu, India.

A.Viswanathan, Lecturer, Department of CSE, K.S.R.College of Engineering, Tiruchengode, Namakkal, Tamilnadu, India

Dr. V. P. Arunachalam, Principal, SNS College of Technology, Coimbatore, Tamilnadu, India

A. Rajiv Kannan, Asst.Prof, Department of CSE, K.S.R.College of Engineering, Tiruchengode, Namakkal, Tamilnadu, India

K. Ganesh Kumar, Lecturer, Department of IT, K.S.R.College of Engineering, Tiruchengode, Namakkal, Tamilnadu, India

Data Group Anonymity: General Approach

Oleg Chertov

Applied Mathematics Department
NTUU “Kyiv Polytechnic Institute”
Kyiv, Ukraine

Dan Tavrov

Applied Mathematics Department
NTUU “Kyiv Polytechnic Institute”
Kyiv, Ukraine

Abstract—In the recent time, the problem of protecting privacy in statistical data before they are published has become a pressing one. Many reliable studies have been accomplished, and loads of solutions have been proposed.

Though, all these researches take into consideration only the problem of protecting individual privacy, i.e., privacy of a single person, household, etc. In our previous articles, we addressed a completely new type of anonymity problems. We introduced a novel kind of anonymity to achieve in statistical data and called it group anonymity.

In this paper, we aim at summarizing and generalizing our previous results, propose a complete mathematical description of how to provide group anonymity, and illustrate it with a couple of real-life examples.

Keywords—group anonymity; microfiles; wavelet transform

I. INTRODUCTION

Throughout mankind’s history, people always collected large amounts of demographical data. Though, until the very recent time, such huge data sets used to be inaccessible for publicity. And what is more, even if some potential intruder got an access to such paper-written data, it would be way too hard for him to analyze them properly!

But, as information technologies develop more, a greater number of specialists (to wide extent) gain access to large statistical datasets to perform various kinds of analysis. For that matter, different data mining systems help to determine data features, patterns, and properties.

As a matter of fact, in today world, in many cases population census datasets (usually referred to as *microfiles*) contain this or that kind of sensitive information about respondents. Disclosing such information can violate a person’s privacy, so convenient precautions should be taken beforehand.

For many years now, mostly every paper in major of providing data anonymity deals with a problem of protecting an individual’s privacy within a statistical dataset. As opposed to it, we have previously introduced a totally new kind of anonymity in a microfile which we called *group anonymity*. In this paper, we aim at gathering and systematizing all our works published in the previous years. Also, we would like to generalize our previous approaches and propose an integrated survey of group anonymity problem.

II. RELATED WORK

A. Individual Anonymity

We understand by *individual data anonymity* a property of information about an individual to be unidentifiable within a dataset.

There exist two basic ways to protect information about a single person. The first one is actually protecting the data in its formal sense, using data encryption, or simply restricting access to them. Of course, this technique is of no interest to statistics and affiliated fields.

The other approach lies in modifying initial microfile data such way that it is still useful for the majority of statistical researches, but is protected enough to conceal any sensitive information about a particular respondent. Methods and algorithms for achieving this are commonly known as *privacy preserving data publishing (PPDP)* techniques. The Free Haven Project [1] provides a very well prepared anonymity bibliography concerning these topics.

In [2], the authors investigated all main methods used in PPDP, and introduced a systematic view of them. In this subsection, we will only slightly characterize the most popular PPDP methods of providing individual data anonymity. These methods are also widely known as *statistical disclosure control (SDC)* techniques.

All SDC methods fall into two categories. They can be either *perturbative* or *non-perturbative*. The first ones achieve data anonymity by introducing some data distortion, whereas the other ones anonymize the data without altering them.

Possibly the simplest perturbative proposition is to add some noise to initial dataset [3]. This is called *data randomization*. If this noise is independent of the values in a microfile, and is relatively small, then it is possible to perform statistical analysis which yields rather close results compared to those ones obtained using initial dataset. Though, this solution is not quite efficient. As it was shown in [4], if there are other sources available aside from our microfile with intersecting information, it will be very possible to violate privacy.

Another option is to reach data *k-anonymity*. The core of this approach is to somehow ensure that all combinations of microfile attribute values are associated with at least *k* respondents. This result can be obtained using various methods [5, 6].

Yet another technique is to *swap* confidential microfile attribute values between different individuals [7].

Non-perturbative SDC methods are mainly represented by *data recoding* (data enlargement) and *data suppression* (removing the data from the original microfile) [6].

In previous years, novel methods evolved, e.g., matrix decomposition [8], or factorization [9]. But, all of them aim at preserving individual privacy only.

B. Group Anonymity

Despite the fact that PPDP field is developing rather rapidly, there exists another, completely different privacy issue which hasn't been studied well enough yet. Speaking more precisely, it is another kind of anonymity to be achieved in a microfile.

We called this kind of anonymity *group anonymity*. The formal definition will be given further on in this paper, but in a way this kind of anonymity aims at protecting such data features and patterns which cannot be determined by analyzing standalone respondents.

The problem of providing group anonymity was initially addressed in [10]. Though, there has not been proposed any feasible solution to it then.

In [11, 12], we presented a rather effective method for solving some particular group anonymity tasks. We showed its main features, and discussed several real-life practical examples.

The most complete survey of group anonymity tasks and their solutions as of time this paper is being written is [13]. There, we tried to gather up all existing works of ours in one place, and also added new examples that reflect interesting peculiarities of our method. Still, [13] lacks a systematized view and reminds more of a collection of separate articles rather than of an integrated study.

That is why in this paper we set a task of embedding all known approaches to solving group anonymity problem into complete and consistent group anonymity theory.

III. FORMAL DEFINITIONS

To start with, let us propose some necessary definitions.

Definition 1. By *microdata* we will understand various data about respondents (which might equally be persons, households, enterprises, and so on).

Definition 2. Respectively, we will consider a *microfile* to be microdata reduced to one file of attributive records concerning each single respondent.

A microfile can be without any complications presented in a matrix form. In such a matrix **M**, each row corresponds to a particular respondent, and each column stands for a specific attribute. The matrix itself is shown in Table I.

TABLE I. MICROFILE DATA IN A MATRIX FORM

		Attributes			
		u_1	u_2	\dots	u_η
Respondents	r_1	ω_{11}	ω_{12}	\dots	$\omega_{1\eta}$
	r_2	ω_{21}	ω_{22}	\dots	$\omega_{2\eta}$
	\dots	\dots	\dots	\dots	\dots
	r_μ	$\omega_{\mu 1}$	$\omega_{\mu 2}$	\dots	$\omega_{\mu \eta}$

In such a matrix, we can define different classes of attributes.

Definition 3. An *identifier* is a microfile attribute which unambiguously determines a certain respondent in a microfile.

From a privacy protection point of view, identifiers are the most security-intensive attributes. The only possible way to prevent privacy violation is to completely eliminate them from a microfile. That is why, we will further on presume that a microfile is always *de-personalized*, i.e., it does not contain any identifiers.

In terms of group anonymity problem, we need to define such attributes whose distribution is of a big privacy concern and has to be thoroughly considered.

Definition 4. We will call an element $s_k^{(v)} \in S_v$, $k = \overline{1, l_v}$, $l_v \leq \mu$, where S_v is a subset of a Cartesian product $u_{v_1} \times u_{v_2} \times \dots \times u_{v_{l_v}}$ (see Table I), a *vital value combination*. Each element of $S_k^{(v)}$ is called a *vital value*. Each u_{v_j} , $j = \overline{1, l_v}$ is called a *vital attribute*.

In other words, vital attributes reflect characteristic properties needed to define a subset of respondents to be protected.

But, it is always convenient to present multidimensional data in a one-dimensional form to simplify its modification. To be able to accomplish that, we have to define yet another class of attributes.

Definition 5. We will call an element $s_k^{(p)} \in S_p$, $k = \overline{1, l_p}$, $l_p \leq \mu$, where S_p is a subset of microfile data elements corresponding to the p^{th} attribute, a *parameter value*. The attribute itself is called a *parameter attribute*.

Parameter values are usually used to somehow arrange microfile data in a particular order. In most cases, resultant data representation contains some sensitive information which is highly recommended to be protected. (We will delve into this problem in the next section.)

Definition 6. A *group* $G(V, P)$ is a set of attributes consisting of several vital attributes $V = \{V_1, V_2, \dots, V_l\}$ and a parameter attribute P , $P \neq V_j$, $j = \overline{1, l}$.

Now, we can formally define a group anonymity task.

Group Anonymity Definition. The task of providing data group anonymity lies in modifying initial dataset for each group $G_i(V_i, P_i)$, $i=1, \dots, k$ such way that sensitive data features become totally confided.

In the next section, we will propose a generic algorithm for providing group anonymity in some most common practical cases.

IV. GENERAL APPROACH TO PROVIDING GROUP ANONYMITY

According to the Group Anonymity Definition, initial dataset \mathbf{M} should be perturbed separately for each group to ensure protecting specific features for each of them.

Before performing any data modifications, it is always necessary to preliminarily define what features of a particular group need to be hidden. So, we need to somehow transform initial matrix into another representation useful for such identification. Besides, this representation should also provide more explicit view of how to modify the microfile to achieve needed group features.

All this leads to the following definitions.

Definition 7. We will understand by a *goal representation* $\Omega(\mathbf{M}, G)$ of a dataset \mathbf{M} with respect to a group G such a dataset (which could be of any dimension) that represents particular features of a group within initial microfile in a way appropriate for providing group anonymity.

We will discuss different forms of goal representations a bit later on in this section.

Having obtained goal representation of a microfile dataset, it is almost always possible to *modify* it such way that security-intensive peculiarities of a dataset become concealed. In this case, it is said we obtain a *modified goal representation* $\Omega'(\mathbf{M}, G)$ of initial dataset \mathbf{M} .

After that, we need to somehow map our modified goal representation to initial dataset resulting in a *modified microdata* \mathbf{M}^* . Of course, it is not necessary that such data modifications lead to any feasible solution. But, as we will discuss it in the next subsections, if to pick specific mappings and data representations, it is possible to provide group anonymity in any microfile.

So, a generic scheme of providing group anonymity is as follows:

1) Construct a (depersonalized) microfile \mathbf{M} representing statistical data to be processed.

2) Define one or several groups $G_i(V_i, P_i)$, $i=1, \dots, k$ representing categories of respondents to be protected.

3) For each i from 1 to k :

a) *Choosing data representation:* Pick a goal representation $\Omega_i(\mathbf{M}, G_i)$ for a group $G_i(V_i, P_i)$.

b) *Performing data mapping:* Define a mapping function $\Upsilon: \mathbf{M} \rightarrow \Omega_i(\mathbf{M}, G_i)$ (called *goal mapping function*) and obtain needed goal representation of a dataset.

c) *Performing goal representation's modification:* Define a functional $\Xi: \Omega_i(\mathbf{M}, G_i) \rightarrow \Omega'_i(\mathbf{M}, G_i)$ (also called *modifying functional*) and obtain a modified goal representation.

d) *Obtaining the modified microfile.* Define an *inverse goal mapping function* $\Upsilon^{-1}: \Omega'_i(\mathbf{M}, G_i) \rightarrow \mathbf{M}^*$ and obtain a modified microfile.

4) *Prepare the modified microfile for publishing.*

Now, let us discuss some of these algorithm steps a bit in detail.

A. Different Ways to Construct a Goal Representation

In general, each particular case demands developing certain data representation models to suit the stated requirements the best way. Although, there are loads of real-life examples where some common models might be applied with a reasonable effect.

In our previous works, we drew a particular attention to one special data goal representation, namely, a *goal signal*. The goal signal is a one-dimensional numerical array $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ representing statistical features of a group. It can consist of values obtained in different ways, but we will defer this discussion for some paragraphs.

In the meantime, let us try to figure out what particular features of a goal signal might turn out to be security-intensive. To be able to do that, we need to consider its graphical representation which we will call a *goal chart*. In [13], we summarized the most important goal chart features and proposed some approaches to modifying them. In order not to repeat ourselves, we will only outline some of them:

1) *Extremums.* In most cases, it is the most sensitive information; we need to transit such extremums from one signal position to another (or, which is also completely convenient, create some new extremums, so that initial ones just “dissolve”).

2) *Statistical features.* Such features as signal mean value and standard deviation might be of a big importance, unless a corresponding parameter attribute is nominal (it will become clear why in a short time).

3) *Frequency spectrum.* This feature might be rather interesting if a goal signal contains some parts repeated cyclically.

Coming from a particular aim to be achieved, one can choose the most suitable modifying functional Ξ to redistribute the goal signal.

Let us understand how a goal signal can be constructed in some widely spread real-life group anonymity problems.

In many cases, we can count up all the respondents in a group with a certain pair of vital value combination and a parameter value, and arrange them in any order proper for a parameter attribute. For instance, if parameter values stand for a person's age, and vital value combinations reflect his or her yearly income, then we will obtain a goal signal representing quantities of people with a certain income distributed by their

age. In some situations, this distribution could lead to unveiling some restricted information, so, a group anonymity problem would evidently arise.

Such a goal signal is called a *quantity signal* $q = (q_1, q_2, \dots, q_m)$. It provides a quantitative statistical distribution of group members from initial microfile.

Though, as it was shown in [12], sometimes absolute quantities do not reflect real situations, because they do not take into account all the information given in a microfile. A much better solution for such cases is to build up a *concentration signal*:

$$c = (c_1, c_2, \dots, c_m) \equiv \left(\frac{q_1}{\rho_1}, \frac{q_2}{\rho_2}, \dots, \frac{q_m}{\rho_m} \right). \quad (1)$$

In (1), ρ_i , $i=1, \dots, m$ stand for the quantities of respondents in a microfile from a group defined by a superset for our vital value combinations. This can be explained on a simple example. Information about people with AIDS distributed by regions of a state can be valid only if it is represented in a relative form. In this case, q_i would stand for a number of ill people in the i^{th} region, whereas ρ_i could possibly stand for the whole number of people in the i^{th} region.

And yet another form of a goal signal comes to light when processing comparative data. A representative example is as follows: if we know concentration signals built separately for young males of military age and young females of the same age, then, maximums in their difference might point at some restricted military bases.

In such cases, we deal with two concentration signals $c^{(1)} = (c_1^{(1)}, c_2^{(1)}, \dots, c_m^{(1)})$ (also called a *main concentration signal*) and $c^{(2)} = (c_1^{(2)}, c_2^{(2)}, \dots, c_m^{(2)})$ (a *subordinate concentration signal*). Then, the goal signal takes a form of a *concentration difference signal* $\delta = (c_1^{(1)} - c_1^{(2)}, c_2^{(1)} - c_2^{(2)}, \dots, c_m^{(1)} - c_m^{(2)})$.

In the next subsection, we will address the problem of picking a suitable modifying functional, and also consider one of its possible forms already successfully applied in our previous papers.

B. Picking Appropriate Modifying Functional

Once again, there can be created way too many unlike modifying functionals, each of them taking into consideration these or those requirements set by a concrete group anonymity problem definition. In this subsection, we will look a bit in detail at two such functionals.

So, let us pay attention to the first goal chart feature stated previously, which is in most cases the feature we would like to protect. Let us discuss the problem of altering extremums in an initial goal chart.

In general, we might perform this operation quite arbitrarily. The particular scheme of such extremums

redistribution would generally depend on the quantity signal nature, sense of parameter values, and correct data interpreting. But, as things usually happen in statistics, we might as well want to guarantee that data utility wouldn't reduce much. By *data utility preserving* we will understand the situation when the modified goal signal yields similar, or even the same, results when performing particular types of statistical (but not exclusively) analysis.

Obviously, altering the goal signal completely off-hand without any additional precautions taken wouldn't be very convenient from the data utility preserving point of view. Hopefully, there exist two quite dissimilar, thought powerful techniques for preserving some goal chart features.

The first one was proposed in [14]. Its main idea is to *normalize* the output signal using such transformation that both mean value and standard deviation of a signal remain stable. Surely, this is not ideal utility preserving. But, the signal obtained this way at least yields the same results when performing basic statistical analysis. So, the formula goes as follows:

$$\theta^* = \left(\theta + \frac{\sigma^*}{\sigma} \cdot \varepsilon - \varepsilon^* \right) \cdot \frac{\sigma}{\sigma^*}. \quad (2)$$

$$\text{In (2), } \varepsilon = \frac{1}{m} \sum_{i=1}^m \theta_i, \varepsilon^* = \frac{1}{m} \sum_{i=1}^m \theta_i^*, \sigma = \sqrt{\frac{\sum_{i=1}^m (\theta_i - \varepsilon)^2}{m-1}},$$

$$\sigma^* = \sqrt{\frac{\sum_{i=1}^m (\theta_i^* - \varepsilon^*)^2}{m-1}}.$$

The second method of modifying the signal was initially proposed in [11], and was later on developed in [12, 13]. Its basic idea lies in applying *wavelet transform* to perturbing the signal, with some slight restrictions necessary for preserving data utility:

$$\theta(t) = \sum_i a_{k,i} \cdot \varphi_{k,i}(t) + \sum_{j=k}^1 \sum_i d_{j,i} \cdot \psi_{j,i}(t). \quad (3)$$

In (3), $\varphi_{k,i}$ stands for shifted and sampled *scaling functions*, and $\psi_{j,i}$ represents shifted and sampled *wavelet functions*. As we showed in our previous researches, we can gain group anonymity by modifying *approximation coefficients* $a_{k,i}$. At the same time, if we don't modify *detail coefficients* $d_{j,i}$ we can preserve signal's frequency characteristics necessary for different kinds of statistical analysis.

More than that, we can always preserve the signal's mean value without any influence on its extremums:

$$\theta_{fin}^* = \theta_{mod}^* \cdot \left(\frac{\sum_{i=1}^m \theta_i}{\sum_{i=1}^m \theta_{mod}^*} \right). \quad (4)$$

In the next section, we will study several real-life practical examples, and will try to provide group anonymity for appropriate datasets. Until then, we won't delve deeper into wavelet transforms theory.

C. The Problem of Minimum Distortion when Applying Inverse Goal Mapping Function

Having obtained modified goal signal θ_{fin}^* , we have no other option but to modify our initial dataset \mathbf{M} , so that its contents correspond to θ_{fin}^* .

It is obvious that, since group anonymity has been provided with respect to only a single respondent group, modifying the dataset \mathbf{M} almost inevitably will lead to introducing some level of data distortion to it. In this subsection, we will try to minimize such distortion by picking sufficient inverse goal mapping functions.

At first, we need some more definitions.

Definition 8. We will call microfile \mathbf{M} attributes *influential ones* if their distribution plays a great role for researchers.

Obviously, vital attributes are influential by definition.

Keeping in mind this definition, let us think over a particular procedure of mapping the modified goal signal θ_{fin}^* to a modified microfile \mathbf{M}^* . The most adequate solution, in our opinion, implies swapping parameter values between pairs of somewhat close respondents. We might interpret this operation as “transiting” respondents between two different groups (which is in fact the case).

But, an evident problem arises. We need to know how to define whether two respondents are “close” or not. This could be done if to measure such closeness using *influential metric* [13]:

$$InfM(r, r^*) = \sum_{p=1}^{n_{ord}} \zeta_p \left(\frac{r(I_p) - r^*(I_p)}{r(I_p) + r^*(I_p)} \right)^2 + \sum_{k=1}^{n_{nom}} \gamma_k \left(\chi(r(J_k), r^*(J_k)) \right)^2. \quad (5)$$

In (5), I_p stands for the p^{th} ordinal influential attribute (making a total of n_{ord}). Respectively, J_k stands for the k^{th} nominal influential attribute (making a total of n_{nom}). Functional $r(\cdot)$ stands for a record's r specified attribute value. Operator $\chi(v_1, v_2)$ is equal to χ_1 if values v_1 and v_2 represent one category, and χ_2 , if it is not so. Coefficients ζ_p and γ_k should be taken coming from importance of a certain attribute (for those ones not to be changed at all they ought to be as big

as possible, and for those ones that are not important they could be zero).

With the help of this metric, it is not too hard to outline the generic strategy of performing inverse data mapping. One needs to search for every pair of respondents yielding minimum influential metric value, and swap corresponding parameter values. This procedure should be carried out until the modified goal signal θ_{fin}^* is completely mapped to \mathbf{M}^* .

This strategy seems to be NP-hard, so, the problem of developing more computationally effective inverse goal mapping functions remains open.

V. SOME PRACTICAL EXAMPLES OF PROVIDING GROUP ANONYMITY

In this subsection, we will discuss two practical examples built upon real data to show the proposed group anonymity providing technique in action.

According to the scheme introduced in Section IV, the first thing to accomplish is to compile a microfile representing the data we would like to work with. For both of our examples, we decided to take 5-Percent Public Use Microdata Sample Files provided by the U.S. Census Bureau [15] concerning the 2000 U.S. census of population and housing microfile data. But, since this dataset is huge, we decided to limit ourselves with analyzing the data on the state of California only.

The next step (once again, we will carry it out the same way for both examples) is to define group(s) to be protected. In this paper, we will follow [11], i.e. we will set a task of protecting military personnel distribution by the places they work at. Such a task has a very important practical meaning. The thing is that extremums in goal signals (both quantity and concentration ones) with a very high probability mark out the sites of military cantonments. In some cases, these cantonments aren't likely to become widely known (especially to some potential adversaries).

So, to complete the second step of our algorithm, we take “Military service” attribute as a vital one. This is a categorical attribute, with integer values ranging from 0 to 4. For our task definition, we decided to take one vital value, namely, “1” which stands for “Active duty”.

But, we also need to pick an appropriate parameter attribute. Since we aim at redistributing military servicemen by different territories, we took “Place of Work Super-PUMA” as a parameter attribute. The values of this categorical attribute represent codes for Californian statistical areas. In order to simplify our problem a bit, we narrowed the set of this attribute's values down to the following ones: 06010, 06020, 06030, 06040, 06060, 06070, 06080, 06090, 06130, 06170, 06200, 06220, 06230, 06409, 06600, and 06700. All these area codes correspond to border, island, and coastal statistical areas.

From this point, we need to make a decision about the goal representation of our microdata. To show peculiarities of different kinds of such representations, we will discuss at least two of them in this section. The first one would be the quantity signal, and the other one would be its concentration analogue.

A. Quantity Group Anonymity Problem

So, having all necessary attributes defined, it is not too hard to count up all the military men in each statistical area, and gather them up in a numerical array sorted in an ascending order by parameter values. In our case, this quantity signal looks as follows:

$$q = (19, 12, 153, 71, 13, 79, 7, 33, 16, 270, 812, 135, 241, 14, 60, 4337).$$

The graphical representation of this signal is presented in Fig. 1a.

As we can clearly see, there is a very huge extremum at the last signal position. So, we need to somehow eliminate it, but simultaneously preserve important signal features. In this example, we will use wavelet transforms to transit extremums to another region, so, according to the previous section, we will be able to preserve high-frequency signal spectrum.

As it was shown in [11], we need to change signal approximation coefficients in order to modify its distribution. To obtain approximation coefficients of any signal, we need to *decompose* it using appropriate *wavelet filters* (both high- and low-frequency ones). We won't explain in details here how to perform all the wavelet transform steps (refer to [12] for details), though, we will consider only those steps which are necessary for completing our task.

So, to decompose the quantity signal q by two levels using Daubechies second-order low-pass wavelet decomposition filter $l \equiv \left(\frac{1-\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{1+\sqrt{3}}{4\sqrt{2}} \right)$, we need to perform the following operations:

$$a_2 = (q *_{\downarrow 2} l) *_{\downarrow 2} l = (2272.128, 136.352, 158.422, 569.098).$$

By $*_{\downarrow 2}$ we denote the operation of convolution of two vectors followed by dyadic downsampling of the output. Also, we present the numerical values with three decimal numbers only due to the limited space of this paper.

By analogue, we can use the flipped version of l (which would be a high-pass wavelet decomposition filter) denoted by $h = \left(\frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{1-\sqrt{3}}{4\sqrt{2}} \right)$ to obtain detail coefficients at level 2:

$$d_2 = (q *_{\downarrow 2} l) *_{\downarrow 2} h = (-508.185, 15.587, 546.921, -315.680).$$

According to the wavelet theory, every numerical array can be presented as the sum of its low-frequency component (at the last decomposition level) and a set of several high-frequency ones at each decomposition level (called *approximation* and *details* respectively). In general, the signal approximation and details can be obtained the following way (we will also substitute the values from our example):

$$A_2 = (a_2 *_{\uparrow 2} l) *_{\uparrow 2} l = (1369.821, 687.286, 244.677, 41.992, -224.980, 11.373, 112.860, 79.481, 82.240, 175.643, 244.757, 289.584, 340.918, 693.698, 965.706, 1156.942);$$

$$D_1 + D_2 = d_1 *_{\uparrow 2} h + (d_2 *_{\uparrow 2} h) *_{\uparrow 2} l = (-1350.821, -675.286, -91.677, 29.008, 237.980, 67.627, -105.860, -46.481, -66.240, 94.357, 567.243, -154.584, -99.918, -679.698, -905.706, 3180.058).$$

To provide group anonymity (or, redistribute signal extremums, which is the same), we need to replace A_2 with another approximation, such that the resultant signal (obtained when being summed up with our details $D_1 + D_2$) becomes different. Moreover, the only values we can try to alter are approximation coefficients.

So, in general, we need to solve a corresponding optimization problem. Knowing the dependence between A_2 and a_2 (which is pretty easy to obtain in our model example), we can set appropriate constraints, and obtain a solution \hat{a}_2 which completely meets our requirements.

For instance, we can set the following constraints:

$$\begin{cases} 0.637 \cdot \hat{a}_2(1) - 0.137 \cdot \hat{a}_2(4) \leq 1369.821; \\ 0.296 \cdot \hat{a}_2(1) + 0.233 \cdot \hat{a}_2(2) - 0.029 \cdot \hat{a}_2(4) \leq 687.286; \\ 0.079 \cdot \hat{a}_2(1) + 0.404 \cdot \hat{a}_2(2) + 0.017 \cdot \hat{a}_2(4) \leq 244.677; \\ -0.137 \cdot \hat{a}_2(1) + 0.637 \cdot \hat{a}_2(2) \geq -224.980; \\ -0.029 \cdot \hat{a}_2(1) + 0.296 \cdot \hat{a}_2(2) + 0.233 \cdot \hat{a}_2(3) \geq 11.373; \\ 0.017 \cdot \hat{a}_2(1) + 0.079 \cdot \hat{a}_2(2) + 0.404 \cdot \hat{a}_2(3) \geq 112.860; \\ -0.012 \cdot \hat{a}_2(2) + 0.512 \cdot \hat{a}_2(3) \geq 79.481; \\ -0.137 \cdot \hat{a}_2(2) + 0.637 \cdot \hat{a}_2(3) \geq 82.240; \\ -0.029 \cdot \hat{a}_2(2) + 0.296 \cdot \hat{a}_2(3) + 0.233 \cdot \hat{a}_2(4) \geq 175.643; \\ 0.233 \cdot \hat{a}_2(1) - 0.029 \cdot \hat{a}_2(3) + 0.296 \cdot \hat{a}_2(4) \geq 693.698; \\ 0.404 \cdot \hat{a}_2(1) + 0.017 \cdot \hat{a}_2(3) + 0.079 \cdot \hat{a}_2(4) \leq 965.706; \\ 0.512 \cdot \hat{a}_2(1) - 0.012 \cdot \hat{a}_2(4) \leq 1156.942. \end{cases}$$

The solution might be as follows: $\hat{a}_2 = (0, 379.097, 31805.084, 5464.854)$.

Now, let us obtain our new approximation \hat{A}_2 , and a new quantity signal \hat{q} :

$$\hat{A}_2 = (\hat{a}_2 *_{\uparrow 2} l) *_{\uparrow 2} l = (-750.103, -70.090, 244.677, 194.196, 241.583, 345.372, 434.049, 507.612, 585.225, 1559.452, 2293.431, 2787.164, 3345.271, 1587.242, 449.819, -66.997);$$

$$\hat{q} = \hat{A}_2 + D_1 + D_2 = (-2100.924, -745.376, 153.000, 223.204, 479.563, 413.000, 328.189, 461.131, 518.985, 1653.809, 2860.674, 2632.580, 3245.352, 907.543, -455.887, 3113.061).$$

Two main problems almost always arise at this stage. As we can see, there are some negative elements in the modified

goal signal. This is completely awkward. A very simple though quite adequate way to overcome this backfire is to add a reasonably big number (2150 in our case) to all signal elements. Obviously, the mean value of the signal will change. After all, these two issues can be solved using the following

$$\text{formula: } q_{mod}^* = (\bar{q} + 2150) \cdot \left(\sum_{i=1}^{16} q_i \right) / \left(\sum_{i=1}^{16} (\bar{q} + 2150) \right).$$

If to round q_{mod}^* (since quantities have to be integers), we obtain the modified goal signal as follows:

$$q_{fin}^* = (6, 183, 300, 310, 343, 334, 323, 341, 348, 496, 654, 624, 704, 399, 221, 686).$$

The graphical representation is available in Fig. 1b.

As we can see, the group anonymity problem at this point has been completely solved: all initial extremums persisted, and some new ones emerged.

The last step of our algorithm (i.e., obtaining new microfile \mathbf{M}^*) cannot be shown in this paper due to evident space limitations.

B. Concentration Group Anonymity Problem

Now, let us take the same dataset we processed before. But, this time we will pick another goal mapping function. We will try to build up a concentration signal.

According to (1), what we need to do first is to define what ρ_i to choose. In our opinion, the whole quantity of males 18 to 70 years of age would suffice.

By completing necessary arithmetic operations, we finally obtain the concentration signal:

$$c = (0.004, 0.002, 0.033, 0.009, 0.002, 0.012, 0.002, 0.007, 0.001, 0.035, 0.058, 0.017, 0.030, 0.003, 0.004, 0.128).$$

The graphical representation can be found in Fig. 2a.

Let us perform all the operations we've accomplished earlier, without any additional explanations (we will reuse notations from the previous subsection):

$$a_2 = (c *_{\downarrow 2} l) *_{\downarrow 2} l = (0.073, 0.023, 0.018, 0.059);$$

$$d_2 = (c *_{\downarrow 2} l) *_{\downarrow 2} h = (0.003, -0.001, 0.036, -0.018);$$

$$A_2 = (a_2 *_{\uparrow 2} l) *_{\uparrow 2} l = (0.038, 0.025, 0.016, 0.011, 0.004, 0.009, 0.010, 0.009, 0.008, 0.019, 0.026, 0.030, 0.035, 0.034, 0.034, 0.037);$$

$$D_1 + D_2 = d_1 *_{\uparrow 2} h + (d_2 *_{\uparrow 2} h) *_{\uparrow 2} l = (-0.034, -0.023, 0.017, -0.002, -0.002, 0.003, -0.009, -0.002, -0.007, 0.016, 0.032, -0.013, -0.005, -0.031, -0.030, 0.091).$$

The constraints for this example might look the following way:

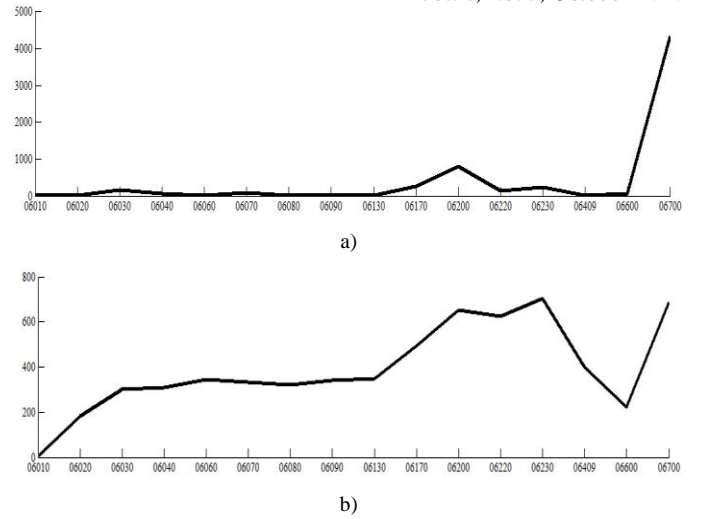


Figure 1. Initial (a) and modified (b) quantity signals.

$$\begin{cases} 0.637 \cdot \hat{a}_2(1) - 0.137 \cdot \hat{a}_2(4) \leq 0.038; \\ 0.296 \cdot \hat{a}_2(1) + 0.233 \cdot \hat{a}_2(2) - 0.029 \cdot \hat{a}_2(4) \leq 0.025; \\ 0.079 \cdot \hat{a}_2(1) + 0.404 \cdot \hat{a}_2(2) + 0.017 \cdot \hat{a}_2(4) \leq 0.016; \\ -0.012 \cdot \hat{a}_2(1) + 0.512 \cdot \hat{a}_2(2) \leq 0.011; \\ -0.137 \cdot \hat{a}_2(1) + 0.637 \cdot \hat{a}_2(2) \geq 0.005; \\ -0.029 \cdot \hat{a}_2(1) + 0.296 \cdot \hat{a}_2(2) + 0.233 \cdot \hat{a}_2(3) \geq 0.009; \\ 0.017 \cdot \hat{a}_2(1) + 0.079 \cdot \hat{a}_2(2) + 0.404 \cdot \hat{a}_2(3) \geq 0.010; \\ -0.012 \cdot \hat{a}_2(2) + 0.512 \cdot \hat{a}_2(3) \geq 0.009; \\ -0.137 \cdot \hat{a}_2(2) + 0.637 \cdot \hat{a}_2(3) \geq 0.009; \\ -0.029 \cdot \hat{a}_2(2) + 0.296 \cdot \hat{a}_2(3) + 0.233 \cdot \hat{a}_2(4) \geq 0.019; \\ 0.233 \cdot \hat{a}_2(1) - 0.029 \cdot \hat{a}_2(3) + 0.296 \cdot \hat{a}_2(4) \leq 0.034; \\ 0.404 \cdot \hat{a}_2(1) + 0.017 \cdot \hat{a}_2(3) + 0.079 \cdot \hat{a}_2(4) \leq 0.034; \\ 0.512 \cdot \hat{a}_2(1) - 0.012 \cdot \hat{a}_2(4) \leq 0.037. \end{cases}$$

One possible solution to this system is as follows: $\hat{a}_2 = (0, 0.002, 0.147, 0.025)$.

We can obtain new approximation and concentration signal:

$$\hat{A}_2 = (\hat{a}_2 *_{\uparrow 2} l) *_{\uparrow 2} l = (-0.003, -0.000, 0.001, 0.001, 0.001, 0.035, 0.059, 0.075, 0.093, 0.049, 0.022, 0.011, -0.004, 0.003, 0.005, 0.000);$$

$$\hat{c} = \hat{A}_2 + D_1 + D_2 = (-0.037, -0.023, 0.018, -0.001, -0.002, 0.038, 0.051, 0.073, 0.086, 0.066, 0.054, -0.002, -0.009, -0.028, -0.026, 0.092).$$

Once again, we need to make our signal non-negative, and fix its mean value. But, it is obvious that the corresponding quantity signal q_{mod}^* will also have a different mean value. Therefore, fixing the mean value can be done in "the quantity domain" (which we won't present here).

Nevertheless, it is possible to make the signal non-negative after all:

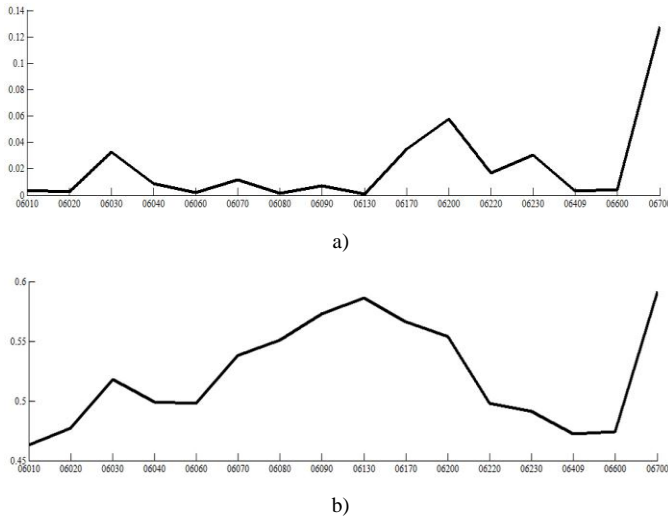


Figure 2. Initial (a) and modified (b) concentration signals.

$c_{mod}^* = \hat{c} + 0.5 = (0.463, 0.477, 0.518, 0.499, 0.498, 0.538, 0.551, 0.573, 0.586, 0.566, 0.554, 0.498, 0.491, 0.472, 0.474, 0.592).$

The graphical representation can be found in Fig. 2b. Once again, the group anonymity has been achieved.

The last step to complete is to construct the modified \mathbf{M}^* , which we will omit in this paper.

VI. SUMMARY

In this paper, it is the first time that group anonymity problem has been thoroughly analyzed and formalized. We presented a generic mathematical model for group anonymity in microfiles, outlined the scheme for providing it in practice, and showed several real-life examples.

As we think, there still remain some unresolved issues, some of them are as follows:

1) *Choosing data representation*: There are still many more ways to pick convenient goal representation of initial data not covered in this paper. They might depend on some problem task definition peculiarities.

2) *Performing goal representation's modification*: It is obvious that the method discussed in Section V is not an exclusive one. There could be as well proposed other sufficient techniques to perform data modifications. For instance, choosing different wavelet bases could lead to yielding different outputs.

3) *Obtaining the modified microfile*: There has to be developed computationally effective heuristics to perform inverse goal mapping.

REFERENCES

- [1] The Free Haven Project [Online]. Available: <http://freehaven.net/anonbib/full/date.html>.
- [2] B. Fung, K. Wang, R. Chen, P. Yu, "Privacy-preserving data publishing: a survey on recent developments," *ACM Computing Surveys*, vol. 42(4), 2010.
- [3] A. Evfimievski, "Randomization in privacy preserving data mining," *ACM SIGKDD Explorations Newsletter*, 4(2), pp. 43-48, 2002.
- [4] H. Kargupta, S. Datta, Q. Wang, K. Sivakumar, "Random data perturbation techniques and privacy preserving data mining", *Knowledge and Information Systems*, 7(4), pp. 387-414, 2005.
- [5] J. Domingo-Ferrer, J. M. Mateo-Sanz, "Practical data-oriented microaggregation for statistical disclosure control," *IEEE Transactions on Knowledge and Data Engineering*, 14(1), pp. 189-201, 2002.
- [6] J. Domingo-Ferrer, "A survey of inference control methods for privacy-preserving data mining," in *Privacy-Preserving Data Mining: Models and Algorithms*, C. C. Aggarwal and P. S. Yu, Eds. New York: Springer, 2008, pp. 53-80.
- [7] S. E. Fienberg, J. McIntyre, *Data Swapping: Variations on a Theme by Dalenius and Reiss*, Technical Report, National Institute of Statistical Sciences, 2003.
- [8] S. Xu, J. Zhang, D. Han, J. Wang, "Singular value decomposition based data distortion strategy for privacy protection," *Knowledge and Information Systems*, 10(3), pp. 383-397, 2006.
- [9] J. Wang, W. J. Zhong, J. Zhang, "NNMF-based factorization techniques for high-accuracy privacy protection on non-negative-valued datasets," in *The 6th IEEE Conference on Data Mining, International Workshop on Privacy Aspects of Data Mining*. Washington: IEEE Computer Society, 2006, pp. 513-517.
- [10] O. Chertov, A. Pilipyuk, "Statistical disclosure control methods for microdata," in *International Symposium on Computing, Communication and Control*. Singapore: IACSIT, 2009, pp. 338-342.
- [11] O. Chertov, D. Tavrov, "Group anonymity," in *IPMU-2010, CCSI*, vol. 81, E. Hüllermeier and R. Kruse, Eds. Heidelberg: Springer, 2010, pp. 592-601.
- [12] O. Chertov, D. Tavrov, "Providing group anonymity using wavelet transform," in *BNCOD 2010, LNCS*, vol. 6121, L. MacKinnon, Ed. Heidelberg: Springer, 2010, in press.
- [13] O. Chertov, *Group Methods of Data Processing*. Raleigh: Lulu.com, 2010.
- [14] L. Liu, J. Wang, J. Zhang, "Wavelet-based data perturbation for simultaneous privacy-preserving and statistics-preserving", in *2008 IEEE International Conference on Data Mining Workshops*. Washington: IEEE Computer Society, 2008, pp. 27-35.
- [15] U.S. Census 2000. 5-Percent Public Use Microdata Sample Files [Online]. Available: <http://www.census.gov/Press-Release/www/2003/PUMS5.html>.

A Role-Oriented Content-based Filtering Approach: Personalized Enterprise Architecture Management Perspective

Imran Ghani, Choon Yeul Lee, Seung Ryul Jeong,
Sung Hyun Juhn
(School of Business IT, Kookmin University, 136-702, Korea)

Mohammad Shafie Bin Abd Latiff
(Faculty of Computer Science and Information Systems Universiti
Teknologi Malaysia, 81310, Malaysia)

Abstract - In the content filtering-based personalized recommender systems, most of the existing approaches concentrate on finding out similarities between users' profiles and product items under the situations where a user usually plays a single role and his/her interests persist identical on long term basis. The existing approaches argue to resolve the issues of cold-start significantly while achieving an adequate level of personalized recommendation accuracy by measuring precision and recall. However, we investigated that the existing approaches have not been significantly applied in the context where a user may play multiple roles in a system simultaneously or may change his/her role overtime in order to navigate the resources in distinct authorized domains. The example of such systems is enterprise architecture management systems, or e-Commerce applications. In the scenario of existing approaches, the users need to create very different profiles (preferences and interests) based on their multiple /changing roles; if not, then their previous information is either lost or not utilized. Consequently, the problem of cold-start appears once again as well as the precision and recall accuracy is affected negatively. In order to resolve this issue, we propose an ontology-driven Domain-based Filtering (DBF) approach focusing on the way users' profiles are obtained and maintained over time. We performed a number of experiments by considering enterprise architecture management aspect and observed that our approach performs better compared with existing content filtering-based techniques.

Keywords: *role-oriented content-based filtering, recommendation, user profile, ontology, enterprise architecture management*

1 INTRODUCTION

The existing content-based filtering approaches (Section 2) claim determining the similarities between user's interests and preferences with product items available in the same category. However, we investigated that these approaches achieve sound results under the situations where a user normally plays a particular role. For instance, in e-Commerce applications a user may upgrade his/her subscription package from normal customer to a premium

customer or vice versa. In this scenario, the existing recommender systems usually manage to recommend the information related to a user's new role. However, if a user wishes the system to recommend him/her products as a premium as well as a normal customer then the user needs to create different profiles (preferences and interests) and has to login based on his/her distinct roles. Likewise, Enterprise Architecture Management Systems (EAMS) emerging from the concept of EA [18] deals with multiple domains whereas a user may perform several roles and responsibilities. For instance, a single user may hold a range of roles such as a planner, analyst and EA managers or a designer and developers or constructors and so on. In addition, a user's role may change over time creating a chain of roles from current to past. This setting naturally leads them to build up very different preferences and interests corresponding to the respective roles. On the other hand, a typical EAMS manages enormous amount of distributed information related to several domains such as application software, project management, system interface design and so on. Each of the domains manages several models, components, schematics, principles, business and technology products or services data, business process and workflow guides. This in turn creates complexity in deriving and managing users' preferences and selecting right information from a tremendous information-base and recommending to the right users' roles. Thus, when the user's role is not specific, the recommendation becomes more difficult in existing content-based filtering techniques. As a result they do not scale well in this broader context. In order to limit the scope, this paper focuses on the scenario of EAMS and the implementation related to e-Commerce systems is left to the future work.

The next section describes a detailed survey of the filtering techniques and their limitations relevant to the concern of this paper.

2 RELATED WORK AND LIMITATIONS

A number of content-based filtering techniques [1][2][3][4][5][10][17] have emerged that are used to personalize information for recommender systems. These techniques are inspired from the approaches used for solving information overload problems [11][15]. As mentioned before in Section 1 that a content-based system filters and recommends an item to a user based upon a description of the item and a profile of the user's interests. While a user profile may either be entered by the user, it is commonly learned from feedback the user provides on items or implicitly obtained from user's recent browsing (RB) activities. The aforementioned techniques and systems usually use data obtained from the RB activities that pose significant limitations on recommendation as summarized in the following table.

TABLE1: LIMITATIONS IN EXISTING APPROACHES

1. There are different approaches to learning a model of the user's interest with content-based recommendation, but no content-based recommendation system can give good recommendations if the content does not contain enough information to distinguish items the user likes from items the user doesn't like in a particular context such as if a user plays different roles in a system simultaneously.
2. The existing approaches do not scale well to filter the information if a user's role is frequently changed which creates a chain of roles (from current to past) for a single user. If the user's role is changed from project manager to EA manager, this leads the users to be restricted to seeing items similar to those not relevant to the current role and preferences.

Based on the above concerns, it has been noted that a number of filtering processing techniques exist which have their own limitations. However, there are no standards to process and filter the data, so we designed our own technique called Domain-based Filtering (DBF).

3 MOTIVATION

Typically, there are three categories of filtering techniques classified in the literature [12] including; (1) ontology based systems; (2) trust network based systems; and (3) context-adaptable systems that consider the current time and place of the user. The scope of this paper, however, is the ontology based systems and we have taken the entire Enterprise

Architecture Management (EAM) area into consideration for ontology-based role-oriented content filtering. This is because of the fact that Enterprise Architecture (EAs) produce vast volumes of models and architecture documents that have actually added difficulties for organization's capability to advance the properties and qualities of its information assets with respect to the user's need. The users need to consult a vast amount the current and previous versions of the EA information assets in many cases to comply with the standards. Though, a number of EAMS have been developed however most of them focus on the content-centric aspect [6][7][8][9] but not on the personalization aspect. Therefore, at EAMS level, there is a need for filtering technique that can select and recommend information which is personalized (relevant and understandable) for a range of enterprise users such as planners, analysts, designers, constructors, information asset owners, administrators, project managers, EA managers, developers and so on to serve for better decision making and information transparency at enterprise-wide level. In order to achieve this feature effectively; the semantics-oriented ontology-based filtering and recommendation techniques can play a vital role. The next section discusses the proposed approach.

4 PHYSICAL AND LOGICAL DOMAINS

In order to illustrate the detailed structure of DBF, it is appropriate to clarify that we have classified two types of domains to deal with the data at EAMS level named physical domains (PDs) and logical domains (LDs). The PDs have been defined to classify enterprise assets knowledge (EAK). The EAK is the metadata about information resources/items including artifacts, models, processes, documents, diagrams and so on using RDFS [14] with class hierarchies (Fig 1) and RDF[13] based triple subject-predicate-object format (Table 2). Basically, the concept of PD is similar to organize the product categories in exiting ontology-based ecommerce systems, such as sales and marketing, project management, data management, software applications, and so on.

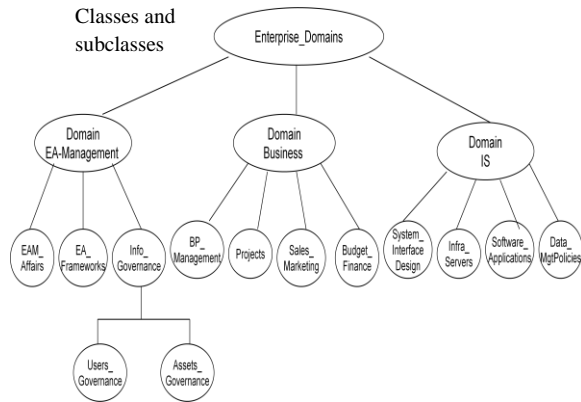


Fig 1: Physical domain (PD) hierarchy

TABLE 2: RDF-BASED INFORMATION ASSETS TRIPLE

Subject (S)	Predicate (P)	Object (O)
EAK_1	dc: title	IT_Governance_Framework
EAK_1	dc: identifier	http://localhost/projects/useam2/sw/EAK/itgov_framework.jpg
EAK_1	useam: Domain	EA_Management
EAK_1	useam: State	Under Development
EAK_1	dc: Description	Aligning IT strategy with the business strategy, Cascading strategy and goals down into the enterprise, Measuring IT's performance
EAK_1	useam: AccessRights	View
EAK_1	dc: creator	Adnan
EAK_1	useam: stdate	2009-01-12
EAK_1	useam: endate	2009-12-31
EAK_1	useam: AccessRoleCode	EA Manager, Software Architect, BP Manager
EAK_1	useam: keyword	IT Governance, IT capability

On the other hand, LDs deal with manipulating the user's profiles organized in user model ontology (UMO). The UMO is organized in Resource Description Framework [13] based triple subject-predicate-object format (Table 3). We name this as LD because it is reconfigurable according to the changing or multiple roles of users and their interests list. Besides, an LD can be deleted if a user leaves the organization. On the other hand PDs are permanent information assets of an enterprise and all the information assets belong to the PDs.

TABLE 3: RDF-BASED UMO

Subject (S)	Predicate (P)	Object (O)
Person	rdf: type	rdfs: Class
foaf: Imran	foaf: name	Imran
foaf: Imran	rdf: type	Person
useam: Imran	useam: hasFctRole	EA_Manager
useam: Imran	useam: hasNfctRole	Not Available
foaf: Imran	useam: hasFctInterest	Reference models, business processes, EA frameworks
foaf: Imran	useam: hasNfctInterest	Not Available
foaf: Imran	foaf: img	Imran.jpg
foaf: Imran	foaf: homepage	resource <http://localhost/useam2/sw/Imran/index.php>

We discuss the DBF approach in the following section.

5 DOMAIN-BASED FILTERING (DBF) APPROACH

As mentioned before in Section 1 that the existing content-base filtering techniques attempt to recommend items similar to those a given user has liked in the past. This mechanism does not scale well in role-oriented settings such as in EAM systems where a user changes his/her role or play multiple roles simultaneously. In this scenario, the existing techniques still bring the old items relevant to the past roles of users which may no longer be desirable to the new role of the user. In our research we worked out to find that there are other criteria that could be used to classify the user's information for filtering purposes. By observing the users' profiles, it has been noted that we can logically distinguish among users' functional and non-functional domains from explicit data collection (when a user is asked to voluntarily provide their valuations including past and current roles and preferences) and implicit data collection (where the user's behavior is monitored) during browsing the system while holding current roles or the roles he/she performed in past.

DBF approach performs its filtering operations by logically classifying the users' profiles based on current and past roles and interests list. Creating LD out of the users' profiles is a system generated process which is achieved by exploring the users' 'roles-interests' similarities as a filtering criterion.

There are two possible criterions to create a user's LD.

- User's current and past roles.
- Users' current and past interests list in accordance with preference-change overtime.

In order to filter and map a user's LD with information in PDs, we have defined two methods.

- Exploring relationships of assets that belong to PD in (EAK) based on LD information (in UMO).

- Information obtained from user's recent browsing (RB) activities. The definition of "recent" may be defined by the organization policy. However, in our prototype we maintain the RB data for one month time.

The working mechanism of our approach is shown in the model below.

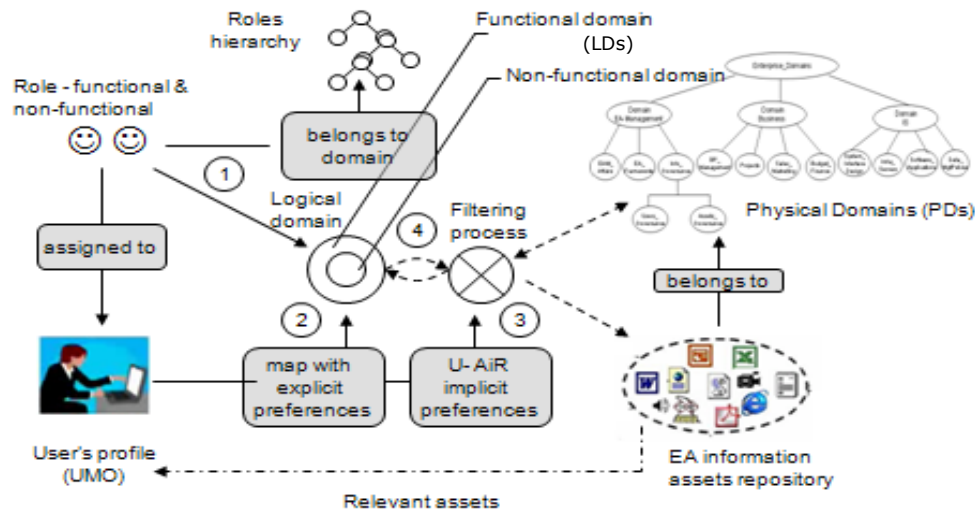


Fig 2: User's relevance with EA information assets based on profile and domain

The Fig 2 is the structure of our model that illustrates the steps to perform role-oriented filtering. At first, we discover and classify the user's functional and non-functional roles and interests from UMO (process 1 and 2 in above figure). As mentioned before, the combination of role and interests list creates the LD of a user. It is appropriate to explain that user's preferred interests are of two types explicit preferences that a user registers in the profile (process 2) and implicit preferences obtained from user's RB activities (process 3). The first type of preference (explicit) is part of UMO that is based on the user's profile while the second type of preferences is part of user-asset information registry (U-AiR) which is a lookup table based on user's RB activity having the potential to be updated frequently. The implicit preferences help to narrow down the results for personalized recommendation level mapping with the most recent interests (in our prototype most recent means one month; however "most recent" period has not been generalized hence

is left to the organizational needs). In our prototype example, our algorithm computes the number of clicks (3~5 clicks) by a user to the concepts on the similar assets (related to the same class or close subclass of the same super class in PDs class hierarchy). If a user performs minimum 3 clicks (threshold) on the concepts of asset then metadata information about that asset is added in to the U-AiR as his/her interested asset assuming that he/she likes that asset. Then, the filtering (process 4) is performed to find the relevant information for the user as shown in the above model (Fig 2). The below Figs 3 (a) (b) illustrate the LD schematic. The outer circle is for functional domain while the inner circle is for non-functional domains. It should be noticed that if user's role is changed to a new role then his/her functional domain is shifted to the inner circle which is for non-functional domain while his old non-functional domain is further pushed downwards. However, the non-functional domain circle may also be overwritten with new non-functional domain depending upon the enterprise strategy. In our prototypical study, we

processed until two levels and keep track the role-change of current and past (recent) record only which is why we have illustrated two circles in Fig 3(a). However, the concept of logical domains is generic thus there may be as many domain-depths (Fig 3(b)) as per the enterprise's policy.

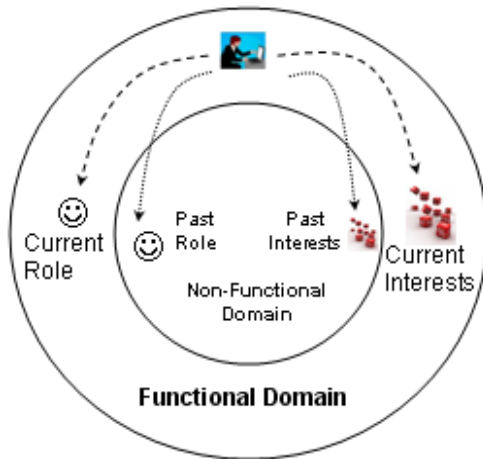


Fig 3 (a): Functional and Non-functional domain schematic

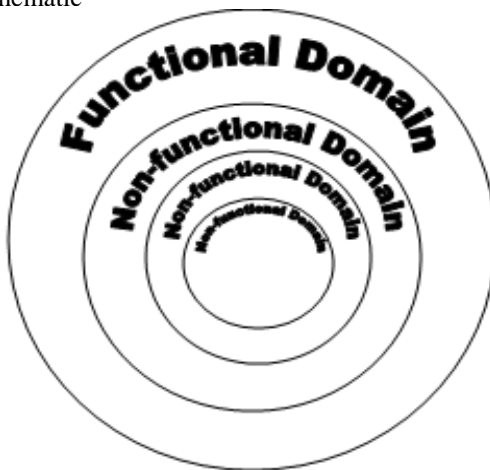


Fig 3 (b): Domains-depth schematic

The next section describes mapping process used for recommendation.

5.1 RELATION-BASED MAPPING PROCESS FOR INFORMATION RECOMMENDATION

In this phase, we traverse the properties of ontologies to find references with roles and EA information assets in domains e.g., sales and marketing, software application and so on. We run the mapping algorithm recursively; for extracting user's attributes information from UMO to create logical functional and non-functional domains and

properties of assets from EAK in order to match the relevance. Three main operations are performed:

(1) The user's explicit profiles are identified in the UMO in the form of concepts, relations, and instances.

- hasFctRole, hasNfctRole, hasFctInterests, hasNfctInterests, hasFctDomain, hasNfctDomain, hasFctCluster, hasNfctCluster, relatesTo, belongTo, conformTo, consultWith, controls, uses, owns, produces and so on

(2) The knowledge about the EA assets is identified

- belongsTo, conformBy, toBeConsulted, consultedBy, toBeControlled, controlledBy, user, owner, and so on

(3) Relationship mapping

The mapping is generated by triggering rules whose conditions match the terms in users' inputs. The user's and information assets attributes are used to formulate rules of the form: IF <condition> THEN <action>, domain= n . The rules indicate the accuracy of the condition that represents the asset in the action part; for example, the information gained by representing document with attributes value 'policy_approval' associated with relation toBeConsulted and belongsTo 'Software_Application'. After acquiring metadata of assets' features, the recommendation system perform classification of user's interests based on the following rules.

Rule1: IF a user Imran's UMO contains predicate 'hasFctRole' which represents the current role and is non-empty with instance value e.g., 'Web programmer' THEN add this predicate with value in functional domain of that user and name it as "ImranFcd" (Imran's functional domain).

Rule2: IF the same user Imran's UMO contains predicate 'hasFctInterests' which represents the current interests and is non-empty with instance value web programming concepts THEN add this predicate the with values in functional domain of that user named as "ImranFcd".

Rule3: IF a user Imran's UMO contains predicate 'hasNFctRole' which represents the past role and is non-empty with instance value e.g., 'EA modeler' THEN add this predicate with value in functional domain of that user and name it as "ImranNFcd".

Rule4: IF the same user Imran's UMO contains predicate 'hasNFctInterests' which represents the past interests and is non-empty with instance value

EA concepts THEN add this predicate with values in functional domain of that user named as “ImranNFcd”.

The process starts by selecting users; keeping in mind domain classification (the system does not relate and recommend information to a user if the domains are not functional or non-functional). The algorithm relates assets items from different classes defined in PDs.

We use the set theory mechanism to match the existing similar concepts. The mapping phase selects concepts from EAK, and maps their attributes with the corresponding concept role in functional and non-functional domains. This mechanism works as a concept explorer, as it detects those concepts that are closely related to the user's roles (functional domain) first and those concepts that are not closely related to the user's roles (non-functional domain) later. In this way, the expected needs are classified by exploring the entities and semantic associations. In order to perform traversals and mapping, we ran the same sequence of instruction to explore the classes and their instances with different parameters of user's LDs and enterprise assets in PDs.

- If node is relevant then continue exploring its properties.
- Otherwise disregard the properties linking the reached node to others in the ontology.

In order to implement the mapping process, we adopted the set theory.

(1)

$$U_i = \{u_{1i}, u_{2i}, u_{3i}, u_{4i}, \dots u_{ni}\}$$

Where i donates the user and ni varies depending on the user's functional roles.

$$D_j = \{u_{1j}, u_{2j}, u_{3j}, u_{4j}, \dots u_{nj}\}$$

Where j donates the assets and n_j varies depending on the assets related to the functional roles.

$$U_i \cap D_j = \alpha_a$$

Where α_a (alpha) is any number of common attributes list.

(2)

$$U_k = \{u_{1k}, u_{2k}, u_{3k}, u_{4k}, \dots u_{nk}\}$$

Where k donates the user and n_k varies depending on the user's non-functional roles.

$$D_k = \{u_{1k}, u_{2k}, u_{3k}, u_{4k}, \dots u_{nk}\}$$

Where k donates the assets and n_k varies depending on the assets related to the non-functional roles.

$$U_k \cap D_k = \alpha_b$$

Where α_b (alpha) is any number of common attributes list.

(3)

$$\alpha_a \cap \alpha_b = \alpha_y$$

Where α_y (alpha) is any number of common attributes list based on the functional and non-functional domains.

A similar series of sets are created for functional and non-functional interest, which are combined to form functional and non-functional domains.

- The stronger the relationship between a node N and the user's profile, the higher the relevance of N.

An information asset, for instance, an article document related to a new EA strategy is relevant if it is semantically associated with at least one role concept in the LDS.

The representation of implementation with scenarios is presented in the next prototypical experiments section.

6 PROTOTYPICAL EXPERIMENTS AND RESULTS

One of the core concerns of an organization is that the people in the organization are performing their roles and responsibilities in accordance with the standards. These standards can be documented in EA [18] and maintained by EAM systems. The EAMSs are used by a number of key role players in the organization including enterprise planners, analysts, designers, constructors, information asset owners, administrators, project managers, EA managers, developers and so on. However, in normal settings to manage and use EA (which is a tremendous strategic asset-base of an organization), a user may perform more than one role such as a user may hold two roles i.e., project manager and EA manager simultaneously. As a result, a user while performing the role as EA manager needs a big-picture top view of all the domains, types of information EA has, EA development process and so on. So, the personalized EAMS should be able to recommend him/her the

information relevant to the big-picture of the organization. On the other hand, if the same user navigates to the project manager role, the EAMS should recommend asset, scheduling policies, information reusability planning among different projects and so on that are specific detailed information relevant to the project domain. Similarly, a user's role may be changed from system interface designer to system analyst. In such a dynamic environment, our DBF approach has the potential to scale well. We implemented our approach in an example job/career service provider company FemaleJobs.Net and conducted evaluations of several aspects of personalization in EAMS prototype. The computed implementation results and users' satisfaction surveys illustrated the viability and suitability of our approach that performed better as compared to the existing approaches at enterprise-level environment. A logical architecture of a personalized EAM is shown in (Fig 4).

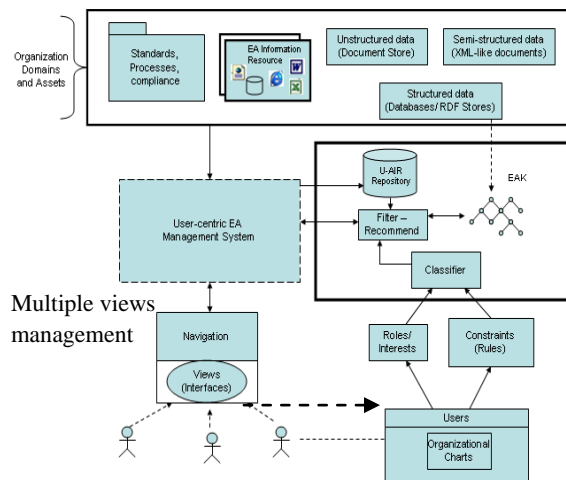


Fig 4: Schematic of personalized EAM system

The above architecture is designed for a web-based tool in order to perform personalized recommendation applicable for EAM and bring the users and EA information assets capabilities together in a unified and logical manner. It interfaces between users and the EA information assets capabilities work together in the enterprise.

Figures 5 and 6 show the browser of personalized information for different types of users based on their functional and non-functional domains.

HOME ABOUT US NEWS YOUR CHOICE SERVICES				
System_InterfaceDesign Domain related artifacts				
Asset Title	Processes Description	Owner/Creator	State	Action
User interface principles	Rules and techniques for generic interfaces for	IT Wing - FemaleJobs.Net	Mature	Recommend
Best practices for User interface	Best practices for generic interfaces for	IT Wing - FemaleJobs.Net	Under_Development	Recommend
Best practices Signup Form	Sign up activities practices for generic	IT Wing - FemaleJobs.Net	Mature	Recommend
Record Removal Policy Alerts	Client side alert for intelligent interfaces	IT Wing - FemaleJobs.Net	Under Development	Recommend

Fig 5: EA information assets recommended to the user based on functional and non-functional domain

HOME ABOUT US NEWS YOUR CHOICE SERVICES				
System_InterfaceDesign Domain related artifacts				
Asset Title	Processes Description	Owner/Creator	State	Action
User interface principles	Rules and techniques for generic interfaces for	IT Wing - FemaleJobs.Net	Mature	Recommend
Best practices for User interface	Best practices for generic interfaces for	IT Wing - FemaleJobs.Net	Under_Development	Recommend
Best practices Signup Form	Sign up activities practices for generic	IT Wing - FemaleJobs.Net	Mature	Recommend
Record Removal Policy Alerts	Client side alert for intelligent interfaces	IT Wing - FemaleJobs.Net	Under Development	Recommend

Fig 6: Interface designer's browser view

We have performed two types evaluations, computational evaluation using precision and recall metrics and anecdotal evaluation using online questionnaire.

6.1 COMPUTATIONAL EVALUATION

The aim of this evaluation was to investigate the role-change occurrences and their impact on users-assets relevance. In this evaluation, we examined whether the highly rated assets are "desirable" to a specific user once his/her role is changed. We compared our approach with existing CMFS[5]. We considered CMFS to perform comparison evaluation with DBF because CMFS pose similarity of editing user's profile with DBF. In DBF, the users' profiles are edited on runtime basis. Besides, the obvious intention was to look into the effectiveness of DBF approach in role-changing environment. In this case even the interest list of a user is populated (based on the user's previous preferences and RB behavior) existing content-based system CMFS was not able to perform filtering operation efficiently. For example, if a user's role is changed the content-based approaches still recommends old items based on the old preferences. The items related to users old role did not appeal to the user, since his responsibilities and preferences were changed. Thus, a user was more interested in new information for compliance of new

business processes. We used precision and recall curve in order to evaluate the performance accuracy for this purpose.

a = the number of relevant EA assets, classified as relevant

b = the number of relevant EA assets, classified as not available.

d = the number of not relevant EA assets, classified as relevant.

Precision = $a/a+d$

Recall = $a/a+b$

We divided this phase into two sub-phases such as before and after the change of user's role respectively.

At first, the user (u_1) was assigned a 'Web programmer' role, and his profile contained explicit interests about web and related programming concepts such as variable and function names conventions, developer manual guide, data dictionary regulations and so on. However, since the user was assigned the role first which is why the implicit interests (likes, dislikes and so on) was not available.

The u_1 started browsing the system. We noted that there were 100 assets in EAK related to u_1 's interests list. We executed the algorithm and computed recall to compare the recommendation accuracy of our DBF approach with existing content-based filtering technique named CMFS.

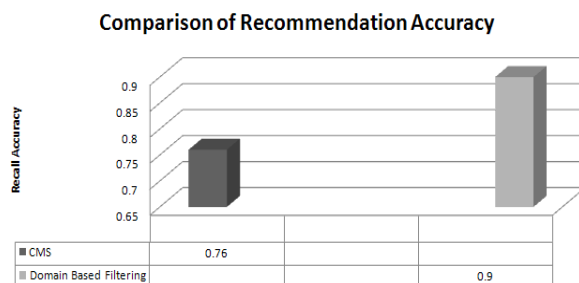


Fig 7: Comparison of CMFS with DBF for recommendation accuracy

The above graph illustrates the comparison analysis showing that our DBF technique significantly performed 18% better than CMS with improved recommendation accuracy measured by recall curve even the sparsness of data [8] was high. This is because of the way the existing techniques

perform filtering based on the explicit and implicit preferences causing cold-start problem [16].

Next, we changed u_1 's role from 'Web programmer' to 'EA Modeler'. After changing the role, user was asked to add explicit concepts regarding new role into the system.

We noted that there were 370 assets related to user's new role. Then, we computed the recommendation accuracy of the approaches using precision after the role-change.

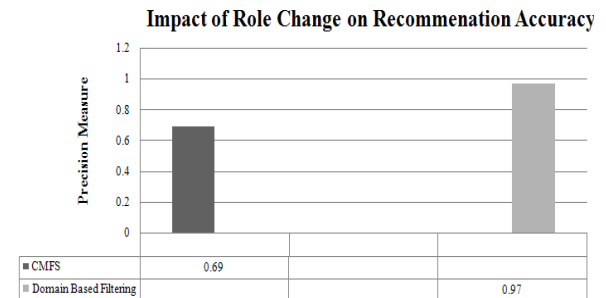


Fig 8: Comparison of approaches for recommendation after role change

As shown in the above two graph (Fig 8) the accuracy of existing technique, after the role-change, reduced by recommending irrelevant assets to user's new role 'EA Modeler' and still bringing up the assets related to the old role 'Web programmer' causing over-specialization again, while, our DBF approach recommend the assets based on the new role because of user's functional and non-functional domain mechanism.

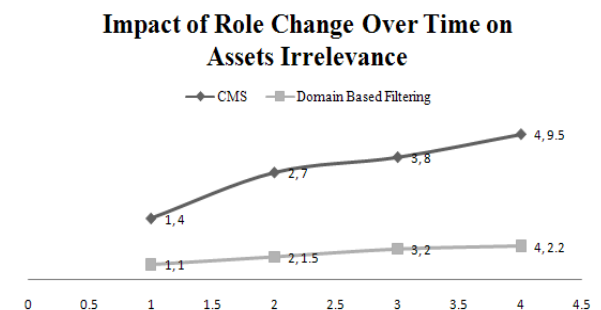


Fig 9: Comparison of approaches for not relevant assets classified as relevant after role change over time

Besides, we also noted the irrelevance of assets while changing the role multiple times. The measurement in Fig 10 shows that DBF approach filtered the EA assets with least irrelevance i.e., 2.2%

compared to CMS with 9.9% irrelevance. This was because of the way existing techniques compute the relevance by considering that the user always performs the same role such as, a customer in e-Commerce website. On the other hand, our DBF approach maintains the user's profiles based on their changing role hence performed better and system recommended more accurately by selecting the assets relevant to the new user's changing role.

6.2 ANECDOTAL EVALUATION

We conducted a survey on the users' experiences about the performance of two EAM platforms i.e., Essentialproject (Fig 10(a)) and our user-centric enterprise architecture management (U-SEAM) system (Fig 10(b)). The survey (Fig 10(c)) was conducted online in an intranet environment. There were two comparison criteria defined for the evaluation. Criteria (1): Personalized information assets aligned with users performing multiple roles simultaneously. Criteria (2) Personalized information assets classified by user's current and past roles.

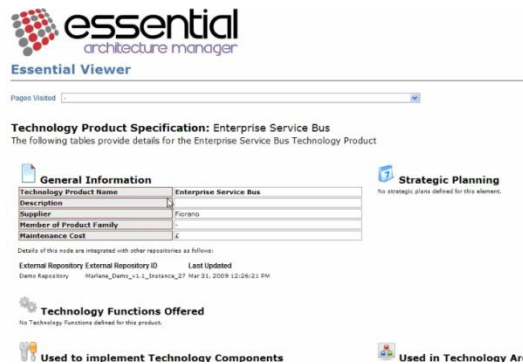


Fig 10(a): Essentialproject EAM System



Fig 10 (b): Our prototype EAMS

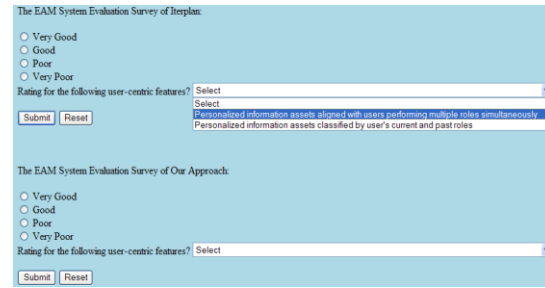


Fig 10 (c): Survey questionnaire

For the evaluation purpose, 12 participants (u1-u12 Fig 11 (a) (b)) were involved in the survey. The users were asked to use both the systems and perform the rating scale as follows: Very Good, Good, Poor and Very Poor. Based on the users' experience, we obtained 144 answers that were used for users' satisfaction analysis. The graphical representation of the user's experience and results can be seen in the following bar charts.

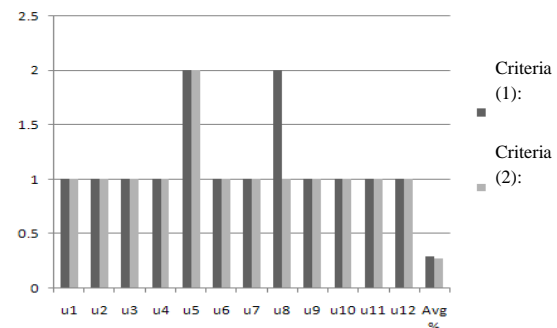


Fig 11 (a): Comparison analysis of EAM systems - Iterplan

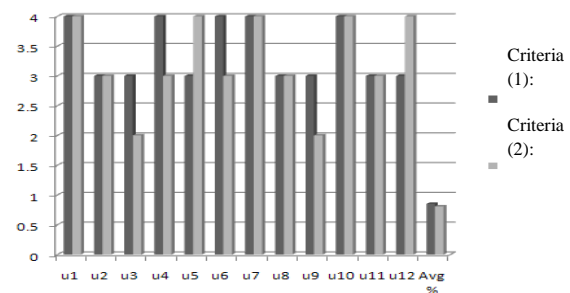


Fig 11 (b): Comparison analysis of EAM systems – Our approach

7 CONCLUSION

We have proposed a novel domain-based filtering (DBF) approach which attempts to increase

accuracy and efficiency of information recommendation. The proposed approach classifies user's profiles in functional and non-functional domain in order to provide personalized recommendation in a role-oriented context that helps improving personalized information recommendation leading towards users' satisfaction.

REFERENCES

- [1] Balabanovic, M.S.Y. (1997). "Fab: content-based, collaborative recommendation." *Communications of the ACM*, v. 40, pp. 66-72.
- [2] Basilico. J and Hofmann. T, A Joint Framework for Collaborative and Content Filtering, In *Proceedings of SIGIR'04*, July 25–29, 2004, Sheffield, South Yorkshire, UK.
- [3] Basu, C.; Hirsh, H. y Cohen, W. (1998). "Recommendation as classification: Using social and content-based information in recommendation." In *Proceedings of the 15th National Conference on Artificial Intelligence*, pp. 714-720.
- [4] Bezerra .B, Carvalho .F. A symbolic approach for content-based information filtering. *Information Processing Letters* Volume 92 , Issue 1 2004. pp: 45 – 52. ISSN:0020-0190.
- [5] Hijikata .Y, Iwahama .K, Takegawa .K, Takegawa .S. Content-based Music Filtering System with Editable User Profile. *SAC'06*, Dijon, France. 2006.
- [6] EEI, Enterprise Elements Inc (2005), *Repository-centric Enterprise-Architecture*. White Paper.
- [7] Essentialproject 2008, <http://www.enterprise-architecture.org>
- [8] Huang. Z, Zeng. D, A Link Analysis Approach to Recommendation under Sparse Data, In *Proceedings of the Tenth Americas Conference on Information Systems*, New York, New York, August 2004.
- [9] iterplan (2009), <http://www.iteraplan.de>
- [10] Kalles. D, Papagelis. A and Zaroliagis. C, Algorithmic aspects of web intelligent systems. In: N. Zhong and Yao,Y. Liu J., Editors, *Web Intelligence*, Springer, Berlin (2003), pp. 323–345.
- [11] Loeb, S. and Terry, D.: *Information Filtering*, *Comm. of the ACM*, Vol. 35, No. 12, pp. 26–81 (1992).
- [12] Peis, E., del Castillo, J.M.M., Delgado-Lopez, J.A. Semantic recommender systems. analysis of the state of the topic. In: *Hipertext*. (2008).
- [13] RDF (1999), www.w3.org/TR/1999/REC-rdf-syntax-19990222
- [14] RDFS (2003), www.w3.org/TR/2003/WD-rdf-schema-20030123.
- [15] Resnick, P. and Varian, H.R.: *Recommender Systems*, *Comm. of the ACM*, Vol. 40, No. 3, pp. 56–89 (1997).
- [16] Schein. A. I, Popescul . A, Ungar. L. H, and Pennock D.M, Methods and metrics for cold-start recommendations. In *ACM SIGIR*, 2002.
- [17] Wang .Y, Stash .N, Aroyo .L, Gorgels .P, Rutledge .P, and Schreiber .G. Recommendations based on semantically-enriched museum collections. *Journal of Web Semantics*, 2008.
- [18] Zachman, J.A., 1987. A framework for information systems architecture. *IBM Syst. J.*, 26(3):276-292.

Minimizing the number of retry attempts in keystroke dynamics through inclusion of error correcting schemes.

Pavaday Narainsamy, Student member IEEE
Computer Science Department,
Faculty of Engineering
University Of Mauritius .

Professor K.M.S.Soyjaudah
Member IEEE
Faculty of Engineering
University of Mauritius

Abstract— One of the most challenging tasks, facing the security expert, remains the correct authentication of human beings. Throughout the evolution of time, this has remained crucial to the fabric of our society. We recognize our friends/enemies by their voice on the phones, by their signature/ writing on a paper, by their face when we encounter them. Police identify thieves by their fingerprint, dead corpse by their dental records and culprits by their deoxyribonucleic acid (DNA) among others. Nowadays with digital devices fully embedded into daily activities, non refutable person identification has taken large scale dimensions. It is used in diverse business sectors including health care, finance, aviation, communication among others. In this paper we investigate the application of correction schemes to the most commonly encountered form of authentication, that is, the knowledge based scheme, when the latter is enhanced with typing rhythms. The preliminary results obtained using this concept in alleviating the retry and account lock problems are detailed.

Keywords— Passwords, Authentication, Keystroke dynamics, errors, N- gram, Minimum edit distance.

I. INTRODUCTION

Although a number of authentication methods exist, the knowledge based scheme has remained the de-facto standard and is likely to remain so for a number years due to its simplicity, ease of use, implementation and its acceptance. Its precision can be adjusted by enforcing password-structure policies or by changing encryption algorithms to achieve desired security level. Passwords represent a cheap and scalable way of validating users, both locally and remotely, to all sorts of services [1, 2]. Unfortunately they inherently suffer deficiencies reflecting from a difficult compromise between security and memorability.

On one hand it should be easy to remember and provide swift authentication. On the other for security purposes it should be difficult to guess, composed of a special combination of characters, changed from time to time, and unique to each account [3]. The larger number and more variability in the set of characters used, the higher is the security provided as it becomes difficult to violate. However such combinations tend to be difficult for end users to remember, particularly when the password does not spell a recognizable word (or includes non-alphanumeric characters such as punctuation marks or other

symbols. Because of these stringent requirements, users adopt unsafe practices such as recording it close to the authentication device, apply same passwords on all accounts or share it with inmates.

To reduce the number of security incidents making the headlines, inclusion of the information contained in the "actions" category has been proposed [4, 5]. An intruder will then have to obtain the password of the user and mimic the typing patterns before being granted access to system resources.

The handwritten signature has its parallel on the keyboard in that the same neuro-physiological factors that account for its uniqueness are also present in a typing pattern as detected in the latencies between two consecutive keystrokes. Keystroke dynamics is also a behavioural biometric that is acquired over time. It measures the manner and the rhythm with which a user types characters on the keyboard. The complexity of the hand and its environment make both typed and written signatures highly characteristics and difficult to imitate. On the computer, it has the advantage of not requiring any additional and costly equipment. From the measured features, the dwell time and flight times are extracted to represent a computer user. The "dwell time" is the amount of time you hold down a particular key while "flight time" is the amount of time it takes to move between keys. A number of commercial products using such schemes already exist on the market [6, 7] while a number of others have been rumored to be ready for release.

Our survey of published work has shown that such implementations have one major constraint in that the typist should not make use of correction keys when keying in the required password. We should acknowledge that errors are common in a number of instances and for a number of reasons. Even when one knows how to write the word, ones fingers may have slipped or one may be typing too fast or pressing keys simultaneously. In brief whatever be the skills and keyboarding techniques used, we do make mistakes, hence the provision for correction keys on all keyboards. Nowadays, typical of word processing softwares, automatic modification based on stored dictionary words can be applied particularly for long sentences. Unfortunately with textual passwords, the text entered is displayed as a string of asterisks and the user cannot spot the

mistake and does make a false login attempt when pressing the enter key. After three such attempts the account is locked and has to be cleared out by the system administrator. Collected figures reveal that between 25% and 50% of help desk calls relate to such problems [8].

Asking the user to input his/her logon credentials all over again instead of using correction keys, clearly demonstrate that inclusion of keystroke dynamics does not seamlessly integrate password mechanism. This can be annoying and stressful for users and will impede on acceptance of the enhanced password mechanism. Moreover this will reduce the probability of the typist correctly matching his enrolled template and hence make another false attempt at login in. In this project we investigate the use of correcting schemes to improve on this limitation and in the long run reduce the number of requests for unlocking account password as encountered by system administrators.

Following this short brief on keystroke dynamics, we'll dwell on the challenges involved in incorporating error correcting techniques technologies to the enhance password mechanism. Our focus will be on a more general approach rather than checking whether the correction keys have been pressed by the user. A scheme that can be customized to deal with cases of damaged keys or American keyboard replaced by English keyboard. In section II, we first review the different correction schemes studied and then the user recognition algorithms to be used before elaborating on an applicable structure for the proposed work. The experimental results are detailed in section V followed by our conclusions and future work in the last section of this paper.

II. BACKGROUND STUDY

To evaluate a biometric system's accuracy, the most commonly adopted metrics are the false rejection rate (FRR) and the false acceptance rate (FAR), which correspond to two popular metrics: sensitivity and specificity [9]. FAR represents the rate at which impostors are accepted in the system as being genuine users while the FRR represents the rate at which authentic users are rejected in the system as they cannot match their template representation. The response of the matching system is a score that quantifies the similarity between the input and the stored representation. Higher score indicates more certainty that the two biometric measurements come from the same person. Increasing the matching score threshold increases the FRR with a decrease in FAR. In practical systems the balance between FAR and FRR dictates the operational point.

A. Error types

Textual passwords are input into systems using keypads/keyboards giving possibilities for typing errors to crop in. The main ones are insertion, deletion, substitution and transposition [10] which amounts to 80 % of all errors encountered [11] with the remaining ones being the split-word and run-on. The last two refer to insertion of space in between characters and deletion of a space between two words respectively. Historically, to overcome mechanical problems associated with the alphabetical order keyboard, the QWERTY layout has been proposed [12] and it has become the de-facto

keyboard used in a number of applications. Other variants exist in "AZERTY" used mainly by French or "QWERTZ" used by Germans. Different keyboarding techniques are adopted by users for feeding data to the device, namely the (i) Hunt and Peck (ii) Touch typing and (iii) Buffering. More information on these can be found in [11]. The first interaction with a keyboard is usually the Hunt and Peck type as the user has to search for the key before hitting on it. Experienced users are considered to be the touch type with a large number of keys being struck per minute.

Typographic errors are due to mechanical failure or slip of the hand or finger, but exclude errors of ignorance. Most involve simple duplication, omission, transposition, or substitution of a small number of characters. The typographic errors for single words have been classified as shown in Table 1 below.

TABLE I. Occurrence of errors in typed text [13]

Errors	% of occurrence
Substitution	40.2
Insertion	33.2
Deletion	21.4
Transposition	5.2

In another work, Grudin [14] investigated the distribution of errors for expert and novice users based on their speed of keying characters. He analysed the error patterns made by six expert typists and eight novice typists after transcribing magazines articles. There were large individual differences in both typing speed and types of errors that were made [15].

The expert users had a range from 0.4% to 0.9% with the majority being insertion errors while for the novice it was 3.2% on average comprising mainly of substitutions ones. These errors are made when the typist knows how to spell the word but may have typed the word hastily. Isolated word error correction includes detecting the error, generating the appropriate candidates for correction and ranking the candidates.

For this project only errors that occur frequently will be given attention as illustrated in table 1 above. Once the errors are detected, they will be corrected through the appropriate correction scheme to enable a legitimate user to log into the system. On the other hand it is primordial that impostors are denied access even though they have correctly guessed the secret code as is normally the case with keystroke dynamics.

B. Error correction

Spell checkers operate on individual words by comparing each of them against the contents of a dictionary. If the word is not found it is considered to be in error and an attempt is made to suggest a word that was likely to have been intended. Six main suggested algorithms for isolated words [16] are listed below.

- 1) The Levenshtein distance or edit distance is the minimum number of elementary editing operations needed to transform an incorrect string of characters into the desired word. The Levenshtein distance caters for three kinds of errors, deletion, insertion and substitution. In addition to its use in spell checkers it has also been applied in speech recognition, deoxyribonucleic acid (DNA) analysis and plagiarism detector [17]. As an example, to transform "symmdtr" to "symmetry" requires a minimum of two operations which are:

- symmdtr → symmetr (substitution of 'd' for 'e')
- symmetr → symmetry (insert 'y' at the end).

Damerau-Levenshtein distance [18] is a variation of the above with the addition of the transposition operation to the basic set. For example to change from 'metirc' to 'metric' requires only a single operation (1 transposition). Another measure is the Jaro-Winkler distance [19] which is a similarity score between two strings and is used in record linkage for duplicate detection. A normalized value of one represents an exact match while zero represents dissimilarity. This distance metric has been found to be best suited for short strings such as people's name [20].

- 2) Similarity key techniques have their strengths in that a string is mapped to a code consisting of its first letter followed by a sequence of three digits, which is same for all similar strings [21]. The Soundex system (patented by Odell and Russell [16, 21]) is an application of such a technique in phonetic spelling correction. Letters are grouped according to their pronunciation e.g. letters "D", "T", "P" and "B" as they produce the same sound. SPEEDCOP (Spelling Error Detection/Correction Project) is a similar work designed to automatically correct spelling errors by finding words similar to the misspelled word [22].
- 3) In rule-based techniques, the knowledge gained from previous spelling error patterns is used to construct heuristics that take advantage of this knowledge. Given that many errors occur due to inversion e.g. the letters *ai* being typed as *ia*, then a rule for this error may be written.
- 4) The N gram technique is used in natural language processing and genetic sequence analysis [23]. An N-gram is a sub-sequence of *n* items (of any size) from a given sequence where the items can be letters, words or base pairs according to the application. In a typed text, unigrams are the single alphabets while digrams (2-gram) are combinations of 2 alphabets taken together.
- 5) The probabilistic technique as the name suggests makes use of probabilities to determine the best correction possible. Once an error is detected, candidate corrections are proposed as different characters are replaced by others using at most one operation. The one having the maximum likelihood is

chosen as the best candidate for the typographical error.

- 6) Neural networks have also been applied as spelling correctors due to their ability to do associative recall based on incomplete and noisy data. They are trained on the spelling errors themselves and once such a scenario is presented they can make the correct inference.

C. Classifier used

Keyboard characteristics are rich in cognitive qualities and as personal identifiers they have been the concern of a number of researchers. The papers surveyed demonstrate a number of approaches that have been used to find adequate keystroke dynamics with a convenient performance to make it practically feasible. Most research efforts related to this type of authentication have focused on improving classifier accuracy [24]. Chronologically it kicked off with statistical classifier more particularly with the T test by Gaines et al [25]. Now the trend is towards the computer extensive neural network variants. Delving into the details of each approach and finding the best classifier to use is well beyond the scope of this project. Our aim is to use one which will measure the similarity between an input keystroke-timing pattern and a reference model of the legitimate user's keystroke dynamics. For that purpose the simple multiple layer perceptron (MLP) with back propagation (BP) used in a previous work was once again considered. A thorough mathematical analysis of the model is presented in the work [26]. It provides details about the why and how of this model. The transfer function used in the neural network was the sigmoid function with ten enrollments for building each user's template.

III. ANALYSIS

The particularity of passwords/secret codes make that they have no specific sound and are independent of any language and may even involve numbers or special characters. Similarity technique is therefore not appropriate as it is based on phonetics and it has limited numbers of possibilities. Moreover with one character and 3 digits for each code there will be frequent collisions as only one thousand combinations exist. Similarly neural network which focuses on the rules of the language for correcting spelling errors turns out to be very complex and inappropriate for such a scenario. A rule based scheme would imply a database of possible errors to be built. Users will have to type a long list of related passwords and best results would be obtained only when the user is making the same error repeatedly. The probabilistic technique uses the maximum likelihood to determine the best correction. The probabilities are calculated from a number of words derived by applying a simple editing operation on the keyed text. Our work involves using only the secret code as the target and the entered text as the input, so only one calculated value is possible, making this scheme useless.

The N-gram technique and the minimum edit distance technique being language and character independent are representative of actual password and were considered for this

project. The distance technique is mostly used for such applications [20].

The N-gram technique compares the source and target words after splitting them into different combination of characters. An intersection and union operations are performed on the different N-grams from which a similarity score is calculated.

Consider two words comprising of eight characters and denoted by source $[s_1 s_2 s_3 s_4 s_5 s_6 s_7 s_8]$ and target $[t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8]$.

2-gram for source: $\{s_1s_2, s_2s_3, s_3s_4, s_4s_5, s_5s_6, s_6s_7, s_7s_8\}$

2-gram for target: $\{t_1t_2, t_2t_3, t_3t_4, t_4t_5, t_5t_6, t_6t_7, t_7t_8\}$

∗: padding space, n(A): number of element in set A.

Union(U) of all digrams = $\{s_1s_2, s_2s_3, s_3s_4, s_4s_5, s_5s_6, s_6s_7, s_7s_8, t_1t_2, t_2t_3, t_3t_4, t_4t_5, t_5t_6, t_6t_7, t_7t_8\}$

Intersection(I) of all digrams = $\{\}$ or null set.

Similarity ratio = $n(I)/n(U)$ ← equation 1

The similarity ratio varies from 0 (which indicates two completely different words) to 1 (words being identical). The process can be repeated for a number of character combinations starting from 2 (di-grams) to the number of characters in the word. From above, if di-grams are considered; for a word length of 8 characters, 1 mistake would give a similarity ratio of 7/11. Seven similar di-grams exist in both words compared to the total set of 11 possible di-graphs with both words taken together.

The Minimum Edit Distance calculates the difference between two strings in terms of number of operations needed to transform one string into another. The algorithm first constructs a matrix with rows being the length of the source word and the column the length of the target word [17]. The matrix is filled with the minimum distance using the operations insertion, deletion and substitution. The last and rightmost value of the matrix gives the minimum edit distance of the horizontal and vertical strings.

The algorithm proceeds as follows.

1. Set n, m to be the length of the source and target words respectively. Construct a matrix containing m rows and n columns.
2. Initialize the first row from 0 to n and the first column from 0 to m incrementally.
3. Consider each character of source (s) (i from 1 to n).
 - a. Examine each character of target (t) (j from 1 to m).
 - Assign cost = 0 to cell value 0 if $s[i]$ equals $t[j]$ else cost = 1.
 - Value allocated to cell is minimum of already filled cells aside + value of 1, i.e upper one ($d[i-1, j]+1$), left one ($d[i, j-1]+1$), c. The cell diagonally above and

to the left plus the cost of current cell ($d[i-1, j-1] + \text{cost cell } (I, j)$).

4. Step 3 is repeated until all characters of source word have been considered.
5. Minimum edit distance is value of the cell (n,m)

IV. SET UP

Capturing keystroke of users is primordial to the proper operation of any keystroke dynamics system. The core of an accurate timing system is the time measuring device implemented either through software or hardware. The latter involves dealing with interrupts, handling processes, registers and addresses which would complicate the design and prevent keystroke dynamics from seamlessly integrating password schemes. Among the different timer options available, the Query Performance Counter (QPC) was used in a normal environment. This approach provided the most appropriate timer for this type of experiment as showed previously [27].

To obtain a reference template, we followed an approach similar to that used by the banks and other financial institutions. A new user goes through a session where he/she provides a number of digital signatures by typing the selected password a number of times. The number of enrollment attempts (ten) was chosen to provide enough data to obtain an accurate estimation of the user mean digital signature as well as information about its variability [28]. Another point worth consideration was preventing annoyance on behalf of the users when keying the same text too many times.

A toolkit was constructed in Microsoft Visual Basic 6.0 which allowed capturing of key depression, key release and key code for each physical key being used. Feature values were then computed from the information in the raw data file to characterize the template vector of each authorized user based on flight and dwell times. One of the issues encountered with efficient typists was release of a key only after s/he has depressed another. The solution was to temporarily store all the key events for a login attempt and then to re-order them so that they were arranged in the order they were first depressed. The typed text collected was then compared to the correct password (string comparison). The similarity score for the N-gram and the minimum edit distance was then computed for the captured text in case no typing mistake was noted, the results being 100%. The user was informed of the presence of inconsistencies noted (use of correction keys) if any when he entered the text. Once accepted the automatic correction was performed and user given access if s/he correctly mimicked his template.

V. RESULTS

The first part of the project was to determine the optimal value of N to be used in the N-gram. The recommended minimum length for password is eight characters [29] and using equation 1, as length increases the similarity score decreases. The number of N-grams in common between the source and target remains the same with different values of N. The total set of possible N-grams increases as length increases. In short for the same error, longer words bring a decrease in the

score. The value of 2 was therefore used. The experiment was performed at university in the laboratory under a controlled environment and users were required to type the text "Thursday" a number of times.

Captured text was then sent to the password correction schemes implemented. Forty users volunteered to participate in the survey and stand as authentic users, the results as computed for users whenever errors are detected is shown below.

TABLE II. Values for each type of error

ERRORS									
Type	Insertion			Substitution			Transposition		
Number	1	2		1	2		1	2	
		C	S		C	S		C	S
Min Edit	1	2	2	1	2		1	2	
N gram	0.75	0.64	0.57	0.67	0.54	0.43	0.54	0.33	0.26

C:Two characters one follow the other.

S:Seperated

They were asked to type their text normally i.e. both with and without errors in their password. They were allowed to use corrections key including the Shift, Insert, Delete and Space bar, Backspace etc. The details were then filtered to get details on those who tried to log into the system as well as the timings for the correct password as entered. Once the threshold for the N gram and Minimum edit was exceeded the system then made the required correction to the captured text. A threshold in the N gram and Min edit controls the numbers of correction keys that can be used. Once the text entered was equivalent to the correct password, the timings were arranged from left to right for each character pressed as it is in the correct password. In case flight time had negative values, they were then arranged in the order they were pressed.

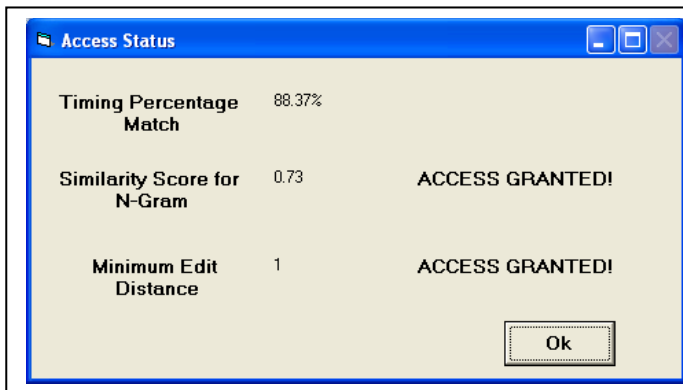


Figure 1: An interaction with the system

Figure 1 above shows an interaction of the user with the system where even with one error in the timing captured the user is being given access to the system.

The possibility for allowing errors in the passwords was then investigated. Though it is not recommended for short words, for long phrase this can be considered such that users do not have to retype the whole text again.

For missing characters the timing used was the one used in the creation of the template but had the highest weight. As reported by Gaines et al [25], each enrollment feature used in building the template is given a weight inversely proportional to its distance from the template value. Accordingly the corrected timing was then sent to the NN toolkit developed as reported in [26].

Out of the 4024 attempts made by all users including impostors, all mistakes using special keys (Insert, Delete, Backspace, Numlock, Space bar) in the typed text could be corrected when it was less than the threshold set (1 error). All genuine users were correctly identified provided they have correctly entered the password and used the correction keys swiftly. Most users who used correction keys and had a considerable increased in the total time taken to key in the password, were not postiviely identified. Moreover those who substituted one character with another and continued typing normally, they were correctly identified.

53 cases remained problematic with the system as there were 2 errors brought into the password. The 2 correction schemes produced results which differed. With a threshold of 0.5 the N gram, did not grant access with 2 transposition mistakes in the password. For 2 errors, the N-gram technique granted the user access while the minimum edit distance technique rejected the user as the threshold was set to 1.

VI. CONCLUSION

Surrogate representations of identity using the password mechanism no longer suffice. In that context a number of studies have proposed techniques which caters for user A sharing his password with user B and the latter being denied access unless he is capable of mimicking the keystroke dynamics of A. Most of the paper surveyed had a major limitation in that when the user makes mistakes and uses backspace or delete key to correct errors, he/she will have to start all over again. In attempt to study the application of errors correcting schemes in the enhanced password mechanism we have we have focused on the commonly used MLP/BP.

TABLE III. Effect of error correction.

	WITHOUT	WITH ERROR CORRECTION
FAR	1%	5%
FRR	8%	15%
REJECTED ATTEMPTS	187	53

The table III above summarizes the results obtained. The FAR which was previously 1% suffered a major degrade in

performance as most users increase their total typing time with the use of correction keys. As expected the FRR changed from 8 to 15 % when errors were allowed in the password. The promising deduction was that using a scheme which allowed one character error in the password, they were correctly identified. Further investigation showed that the major hurdle is with the use of correction keys their normal flow of typing is disrupted and produces false results in keystroke dynamics. This clearly demonstrates the possibility for authenticating genuine users even when the latter has made errors. We have investigated the use of N-gram and minimum distance as they can be varied to check for any error or even allow a minimum of errors to be made. For the latter, with transposition and insertion errors the timings captured could easily cater for the correct password. The main issue encountered was to find a convenient scheme to replace the missing ones. We have adapted our work with the one documented in Gaines et al [25], where we assume that the attempts closest to the template is more representative of the user. The results obtained demonstrate the feasibility of this approach and will boost further research in that direction.

Our focus has been on commonly encountered errors but other possibilities include the use of run on and split word errors among others. Other works that can be carried out along that same line include the use of adaptive learning to be more representative of the user. Logically this will vary considerably as users get acquainted to the input device. Similarly investigation on the best classifier to use with this scheme remains an avenue to explore. An intruder detection unit placed before the Neutral Neural network can enhance its usability and acceptability as a classifier. By removing the intruder attempts and presenting only authentic users to the neutral network an ideal system can be achieved even with learning sample consisting of fewer attempts.

ACKNOWLEDGMENT

The authors are grateful to the staff and students who have willingly participated in our experiment. Thanks extended to those who in one way or another have contributed to make this study feasible.

REFERENCES

- [1] CP, Pfleeger (1997), "Security in Computing", International Edition, Second Edition, Prentice Hall International, Inc,
- [2] S. Garfinkel & E. H. Spafford (1996), "Practical UNIX Security", O Reilly, 2nd edition, April 1996.
- [3] S. Wiedenbeck, J. Waters, J. Birget, A. Brodskiy & Nasir Memon (2005), "Passpoints: Design and Longitudinal Evaluation of a Graphical Password System", International Journal of Human-Computer Studies, vol 63(1-2), pp. 102-127.
- [4] A. Mészáros, Z. Bankó, L. Czúni(2007), "Strengthening Passwords by Keystroke Dynamics", IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, Dortmund, Germany.
- [5] D. Chuda & M. Durfina(2009), "Multifactor authentication based on keystroke dynamics", ACM International Conference Proceeding Series; Vol. 433, Proceedings of the International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing, Article No.: 89
- [6] www.biopassword.com
- [7] www.psylock.com
- [8] I. Armstrong (2003) "Passwords exposed: Users are the weakest link", SCMag June 2003 . Available:<http://www.scmagazine.com>
- [9] S. Y. Kung, M. W. Mak, S. H. Lin(2005), "Biometric Authentication", New Jersey: Prentice Hall, 2005.
- [10] Loghman Barari, Behrang QasemiZadeh(2005),"CloniZER Spell Checker, Adaptive, Language Independent Spell Checker",Proc. of the first ICGST International Conference on Artificial Intelligence and Machine Learning AIML 05, pp 66-71.
- [11] Wikipedia, "Typing" . Available: <http://en.wikipedia.org/wiki/Typing> .
- [12] Sholes, C. Latham; Carlos Glidden & Samuel W. Soule (1868), "Improvement in Type-writing Machines", US 79868, issued July 14.
- [13] James Clawson, Alex Rudnick, Kent Lyons, Thad Starner(2007) ,"Automatic Whiteout: Discovery and Correction of Typographical Errors in Mobile Text Input", Proceedings of the 9th conference on Human-computer interaction with mobile devices and services, New York, NY, USA, 2007. ACM Press. available at http://hackmode.org/~alex/pubs/automatic-whiteout_mobileHCI07.pdf.
- [14] Grudin, J.T. (1983), "Error Patterns in Novice and Skilled Transcription Typing". In Cognitive Aspects of Skilled Typewriting. Cooper, W.E. (ed.). Springer Verlag. ISBN 0-387-90774-2.
- [15] Kukich, K. (1992), "Automatic spelling correction: Detection, correction and context-dependant techniques". Technical report, Bellcore, Morristown, NJ 07960.
- [16] Michael Gilleland(2001), "Levenshtein Distance, in Three Flavors", Available: <http://www.merriampark.com/ld.htm>.
- [17] Wikipedia,"Damerau-Levenshtein distance", Available:http://en.wikipedia.org/wiki/Damerau%E2%80%93Levenshtein_distance
- [18] Winkler, W. E. (1999). "The state of record linkage and current research problems". *Statistics of Income Division, Internal Revenue Service Publication,R99/0*, Available: <http://www.census.gov/srd/papers/pdf/rr99-04.pdf>.
- [19] Wikipedia,"Jaro-Winkle distance". Available: <http://en.wikipedia.org/wiki/Jaro-Winkle>
- [20] DominicJohnRepici(2002), "Understanding Classic SoundEx Algorithms".Available-at <http://www.creativyst.com/Doc/Articles/SoundEx1/SoundEx1.htm>
- [21] Pollock, J. J., and Zamora, A. (1984). "Automatic spelling correction in scientific and scholarly text". Communications of the ACM, 27(4), pp 358-368(1984).
- [22] Wikipedia,"N-gram". Available: <http://en.wikipedia.org/wiki/N-gram>.
- [23] S. Cho & S. Hwang(2006), "Artificial Rhythms and Cues for Keystroke Dynamics Based Authentication", D. Zhang and A.K. Jain (Eds.): Springer-Verlag Berlin Heidelberg , ICB 2006, LNCS 3832, pp. 626 – 632.
- [24] R. Gaines et al (1980), "Authentication by Keystroke Timing: Some Preliminary Results", technical report R-256-NSF, RAND.
- [25] Pavaday N & Soyjaudah. K.M.S, "Investigating performance of neural networks in authentication using keystroke dynamics", In Proceedings of the IEEE africon conference , pp. 1 – 8, 2007.
- [26] Pavaday N , Soyjaudah S & Mr Nugessur Shrikaant, "Investigating & improving the reliability and repeatability of keystroke dynamics timers", International Journal of Network Security & Its Applications (IJNSA), Vol.2, No.3, July 2010.
- [27] Revett, K., Gorunescu, F., Gorunescu, M., Ene, M., de Magalhães, S.T. and Santos, H.M.D. (2007) "A machine learning approach to keystroke dynamics based user authentication", International Journal of Electronic Security and Digital Forensics, Vol. 1, No. 1, pp.55–70.
- [28] Patricia A Wittich (2003), "Biometrics: Are You Key to Security?", pp 1 – 12 , SANS Institute 2003.
- [29] Sebastian Deorowicz, Marcin G.Ciura(2005),"Correcting the spelling errors by modelling their causes", International Journal of Applied. Mathematics. Computer. Science., 2005, Vol. 15, No.2,pp 275–285.

AUTHORS PROFILE

Mr. N. Pavaday is now with the Computer Science, Faculty on Engineering, University of Mauritius, having previously done his research training with the Biometric Lab, School of Industrial Technology, University of Purdue West Lafayette, Indiana, 47906 USA, (phone: +230-4037727 e-mail: n.pavaday@uom.ac.mu).

Professor K.M.S.Soyjaudah is with the same university as the first author. He is interested in all aspect of communication with focus on improving its security. He can also be contacted on the phone +230 403-7866 ext 1367 (e-mail: ssoyjaudah@uom.ac.mu)

Development of Cinema Ontology: A Conceptual and Context Approach

Dr. Sunitha Abburu
Professor & Director
Department of Computer Applications
Adhiyamaan College of Engineering
Hosur, India.
918050594248

Jinesh V N
Lecturer, Department of Computer Science
The Oxford College of Science
Bangalore, India.
919739072949

Abstract— Stored multimedia data poses a number of challenges in the management of multimedia information, includes knowledge representation, indexing and retrieval, intelligent searching techniques, information browsing and query processing. Among the multimedia entertainment, cinema stands in the first position. Ontology is a kind of concept model that could describe system at the level of semantic knowledge as agreed by a community of people. Ontology is hierarchical and thus provides a taxonomy of concepts that allows for the semantic indexing and retrieval of information. Ontology together with a set of individual instances of classes constitutes a knowledge base. In an abstract sense, we view cinema ontology as a collection of sub ontologies. Most of the queries are based on two different aspects of the multimedia objects pertaining to cinema domain viz context information and concept based scenes. There is a need for two kinds of sub ontology pertaining to the cinema domain. Cinema Context Ontology and Cinema Scene Ontology. The former deals with the external information and while the later focus on the semantic concepts of the cinema scene and their hierarchy and the relationship among the concepts. Further practical implementation of Cinema ontology is illustrated using the protégé tool. Finally, designing and construction of context information extraction system and cinema scene search engine are proposed as future work. The proposed structure is flexible and can be easily enhance.

Keywords- *Domain ontology; Concept; Context; Cinema; Multimedia*

I. INTRODUCTION

In this busy and competitive world entertainment media plays a vital role. All need some kind of entertainment to come out of the daily life pressure. The volume of digital video has grown tremendously in recent years, due to low cost digital cameras, scanners, and storage and transmission devices. Multimedia objects are now employed in different areas such as entertainment, advertising, distance learning, tourism, distributed CAD/CAM, GIS, sports etc. This trend has resulted in the emergence of numerous multimedia repositories that require efficient storage. The stored multimedia data poses a number of challenges in the management of multimedia

information, including data and knowledge representation, indexing and retrieval, intelligent searching techniques, information browsing and query processing. Among the multimedia entertainment cinema stands in the first position. Large numbers of groups are involved in the cinema domain. Nowadays multimedia entertainment became more and more popular and vast numbers of groups are working on this domain. Most of entertainment media is introducing cinema related programs. Today's busy world most of us prefer to watch favorite scenes. Our studies on user requirements pertaining to the entertaining of cinema lovers would like to watch information about cinema celebrities like date of birth, hobbies, list of flopped cinemas, ranking...etc., And would also like to view scenes pertaining to specific theme, actor ...etc, they may be looking for their favorite actor, director, musician...etc., At the same time directors, cameramen, stunt masters...etc, would like to view scenes pertaining to a specific theme or different theme to improve or enhance their capabilities, skills or knowledge. Cinema, clipping and related information is/are available in the internet. To improve the effectiveness and efficiency of system, one must concentrate on user community and their requirements in different aspects.

Multimedia objects are required for variety of reasons in different contexts. Video data is rapidly growing and playing a vital role in our life. Despite the vast growth of multimedia objects and information, the effectiveness of its usage is very limited due to the lack of complete organized knowledge representation. The Domain Knowledge should be extracted and stored in an organized manner which will support effective retrieval system. An ontology defines a common vocabulary, common understanding of the structure of domain knowledge among the people who needs to share information. The use of ontology in information systems provides several benefits like knowledge needed and acquired can be stored in a standardized format that unambiguously describes the knowledge in a formal model. Ontology is hierarchical and thus provides a taxonomy of concepts that allows for the semantic indexing and retrieval of information. Ontology

provides a means of data fusion by supplying synonyms or concepts defined using various descriptions. Above points shows the need of cinema ontology.

The rest of the paper is organized as follows. The literature survey report is in section 2. Section 3 discusses the proposed method for cinema domain ontology construction. In section 4, we present a practical implementation and experimental results. Finally, we conclude with a summary and some directions of future research in section 5.

II. LITERATURE ON ONTOLOGY

Ontology has been developed in the artificial intelligence community to describe a variety of domains, and has been suggested as a mechanism to provide applications with domain knowledge and to facilitate the sharing of information [1] [2] [3] [4]. Ontology is a formal, explicit specification of a shared conceptualization [5]. A conceptualization of some phenomenon in the world identifies and determines the relevant concepts and the relations of that phenomenon. Ontology is typically defined as an abstract model of a domain of interest with a formal semantics in the sense that they constitute a logical theory. These models are supposed to represent a shared conceptualization of a domain as they are assumed to reflect the agreement of a certain community or group of people. In the simplest case, ontology consist of a set of concepts or classes which are relevant for the domain of interest as well as a set of relations defined on these concepts. Ontology is a kind of concept model that could describe system at the level of semantic knowledge as agreed by a community of people. It serves as semantic reference for users or applications that accept to align their interpretation of the semantics of their data to the interpretation stored in the ontology [6]. As a new kind of knowledge organization tool, ontology has attracted more and more attention.

Ontology has been widely used in many fields, such as knowledge representation, knowledge sharing, knowledge integration, knowledge reuse, information retrieval, and so on. Hence the development of ontology is seriously impeded [5]. In the field of knowledge engineering, different scholars give different definitions of ontology according to the content of ontology, the form of ontology or the purpose of ontology [7]. Different types of ontology may exist, ranging from sophisticated dictionaries to rich conceptual and formal descriptions of concepts with their relationships and constraints. N. F. Noy, and D. L. McGuinness in [8] describe the need for ontology as:

- To share common understanding of the structure of information among people or software agents.
- To enable reuse of domain knowledge.
- To make domain assumptions explicit.
- To separate domain knowledge from the operational knowledge.
- To analyze domain knowledge.

Ontology development process is an iterative process that will continue in the entire life cycle of the Ontology. The basic steps for building Ontology are:

- Determine the domain and scope of the ontology.
- Consider reusing existing ontology.
- Enumerate important terms in the ontology.
- Define the classes and the class hierarchy.
- Define the properties of classes—slots.
- Define the facets of the slots.
- Create instances.

When ontology is applied to specific field, it refers as domain ontology and is the specification of a particular domain conceptualization. Ontology together with a set of individual instances of classes constitutes a knowledge base. In reality, there is a fine line where the ontology ends and the knowledge base begins. Lili Zhao and Chunping Li [9] proposed ontology based mining for movie reviews which uses the ontology structure as an essential part of the feature extraction process by taking relationship between concepts. The author is using two models movie model and feature model. Amancio Bouza [10] initiated a project on movie ontology with the aim to standardize the representation of movies and movie attributes across data bases. This project provides a controlled vocabulary of semantic concepts and the semantic relations among those concepts. But still, this ontology needs further investigation in collecting, filtering and normalizing concepts, properties, and instances. Shiyao Ou, et al [11] presents an automatic question pattern generation for ontology-based question answering for cinema domain. We have chosen movie domain for the same reasons given by Gijis Geleijnse[12].Gijis have chosen to study the movie domain for two reasons, firstly, numerous web pages handle this topic, the query 'movie' in Google results in 180,000,000 hits. The performance of the algorithm will thus not or barely be influenced by the lack of data available. Secondly, we can easily verify the results and formulate benchmarks for evaluation purposes. To the best of our knowledge, the need and construction of cinema domain ontology has almost not been dealt with. In this paper we present a novel solution to construct cinema domain ontology.

III. CINEMA DOMAIN ONTOLOGY

Cinema domain ontology contains concepts, relations between concepts, concepts attributes. The concept attributes share object oriented structure. Cinema industry involves heterogeneous systems and people. This is the biggest industry in the entertainment world and more complex. As more number of people with different technical, skills and background are trying to show their skills in to the cinema industry. People from vast and various fields are competing to show case their talents, knowledge and the skill sets. All the communities would be interested to know, acquire the knowledge of the latest and best techniques, styles in their

own fields. So our proposed Cinema Ontology model supports such kind of requirement.

In an abstract view, we view cinema ontology as a collection of sub ontologies. The proposed structure is flexible and can be easily enhanced. We represent the cinema ontology (CO) as a collection of sub ontology. $CO = \{CCo, CSo, CMo \dots\}$ Domain Knowledge is required to capture the metadata and annotation in different aspects, as well as to interpret the query. Multimedia objects are required for variety of reasons in different contexts by different communities. We derive the word stakeholder from software engineering aspect for cinema domain as anyone or groups, who are involved, associated or interested in the cinema industry. It is sensible to look for natural subgroups that will be more homogeneous than the total population. Hence in our case we classified the stakeholder community into two classes based on their roles they perform with respect to the cinema domain. Stakeholders who are involved and associated are fall in one class and the interested will fall in other class. The advantage of such classification is that we can easily sum up the retrieval behavior which directly conveys the information requirement. End user's information requirement is a very significant and substantial input during database design. Unfortunately this input will not be readily available and has to be manually collected and accumulated from the real world. Thus it involves extensive human expertise and experience. Even after accumulation there is no guaranty that the information is complete and correct. This has motivated us to design a cinema domain ontology which is flexible and easy to enhance as and when the requirements changes.

As per our literature survey not much work has been done in cinema domain. Survey on stake holders information requirements pertaining to the cinema lovers reflects that they would like to watch information about cinema celebrities like date of birth, hobbies, list of flopped cinemas, ranking...etc, and also would like to view scenes related to specific theme, actor, director, musician...etc,. Where as directors, cameraman, stunt masters, and technical groups...etc, would like to view scenes pertaining to a specific theme or different theme to improve or enhance their capabilities, skills or knowledge. Themes may based on the interest of the viewer pertaining to

- Actor, comedian...
- Actions. (Happy, angry...etc,.)
- Events (celebrations-birthday, wedding, inauguration)
- Location (hill stations, 7 wonders of the world, etc.)
- Settings used.
- Subjective emotions (happiness, violence)
- Costumes used...etc, (dragon, devil, god, special characters...)
- Including the presence of a specific type of object (trains, cars,etc,..)

Metadata such as who captured the video, where, when...etc, motivated by these demands efforts have been made to build a semantic cinema ontology , exploring more efficient context management and information retrieval system. Our studies on user requirements concluded that most of the queries are based on the two different aspects of the multimedia objects viz context, concept based scenes. As the requirements may be related to the cinema scene or the information about the cinema. Cinema domain ontology is a hierarchical structured set of concepts describing cinema context, cinema scenes domain knowledge, which can support cinema information extraction, storage and retrieval system. This gives the need of two kinds of sub ontology pertaining to the cinema domain.

- Cinema context ontology(CCo)
- Cinema Scene ontology (CSo).

In this scenario, the formalism of knowledge must be convenient for structuring the movie descriptions based on available resource.

A. Cinema Context Ontology

The cinema is closely associates with different kinds of information like cinema, cinema celebrities, banner, cinema ranking, etc. This kind of information or data is not related to the content or semantics of the cinema. To represent the complete cinema domain knowledge semantic information must be associated with context information along with the cinema scenes. Moreover the index considering only semantics ignores the context information regarding the video. Unfortunately cinema scenes or multimedia object which is separated from its context has less capability of conveying semantics. For example, diagnostic medical videos are retrieved not only in terms of video content but also in terms of other information associated with the video (like physician's diagnosis, physician details, treatment plan, photograph taken on. ...etc.,). Context information includes information regarding the cinema, such as date of release, place of release, director, producer, actors and so on. In the cinema domain context information abstracts complete information of that context i.e., actors, producers, technical community personal details ...etc,. The context information associated to the cinema domain can be classified in to context independent information and context dependent information as shown in Fig. 1

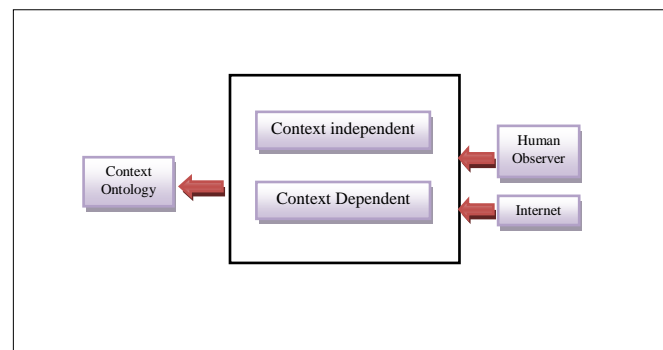


Figure.1 Context Information Classification

Context Dependent Information: The information associated with a particular cinema, like actors, directors' performance in a specific cinema, group of people who are involved in a cinema, team's performance in a cinema i.e. all information associated to a particular cinema.

Context Independent Information: The general details about cinema celebrities like personal details, location details, movie hall details...etc means the information which does not depends upon a particular cinema.

The stake holder would like get the information about the cinema and cinema celebrities. This gives the need for cinema context sub ontology. Ontology plays a more important role in design and sharing of information. To make full use of available data and more efficient search for desired information we need proper representation of knowledge. Effective structure of the knowledge improves the efficiency of the retrieval system. In cinema context ontology the knowledge is represented in such a way that the extraction or retrieval process will be improved and it will be based on context in the cinema. Context of the cinema like actors, director, producer, story writer, editor, cameraman, banner, release date ,success rate, awards won by etc, are is purely text data which can be dealt as information extraction ,storage and retrieval. To support these activities and to improve the efficiency of the retrieval system information is stored and retrieved based on the context sub ontology.

A. Cinema Scene Ontology

Domain ontology is greatly useful in knowledge acquisition, sharing and analysis. In order to acquire the richness and the entertainment contained in cinema we are introducing cinema ontology. Craze on cinema celebrates and cinema scene acted, directed, edited ...etc, by specific individuals of cinema industry ...etc, are too high. The current work is targeted for the cinema stakeholders. The stake holder would like to watch the scenes of cinema celebrities. This gives the need for the cinema scene database, where the main repository is cinema scenes from various movies based on the user interest. For all the above reasons there is a need to define cinema scene ontology. The semantic concepts in generic to cinema domain concept hierarchy and relationship between the concepts, attributes of the concepts ...etc, needs to be considered. To support the cinema scene retrieval, there is a need for cinema scene ontology in which knowledge pertaining to cinema scenes can be identified, represented and classified. The concepts of cinema scenes are identified and classified in general by considering multiple cinemas. Themes, actors, actions, location, action ...etc, are different concepts of the cinema scene ontology.

Video scenes can be queried to their semantic content which will increase the retrieval efficiency. The cinema scene sub ontology supports semantic and concept based retrieval of cinema scenes. CO can reflect the cinema domain knowledge

hierarchical structure clearly and can be used for application systems such as video search engines for wide range of cinema audio, producers, cameramen, directors, musicians ...etc,. Cinema scene sub ontology is based on the cinema scenes and classification of cinema scenes which may help various group of people involved in cinema industry and the TV channels to go for theme, actor oriented entertainment programs. In this ontology the data is arranged in such a way that the extraction or retrieval process will be improved and it will be based on scene in the cinema.

IV. A PRACTICAL APPROACH

Sousan W.L, Wylie, K.L, Zhengxin Chen in [13] describes the method to construct Ontology from text. Xinli and Zhao in [14] studied the government ontology and in [15] construction of university course ontology is discussed. This section enters in to details of the method of constructing cinema Ontology. Row cinema is observed by the human observer and is segmented into scenes. Based on the theme in the scene ontology these scenes are stored in a scene database. Each scene in the scene database is again given to the human observer to identify various concepts instances of the scene, to create the annotation. The scene annotation supports the cinema scene search and retrieval based on various concepts like themes, actor, location, action, event ...etc, is as shown in Fig.2. Context dependent and context independent details are extracted and stored using object oriented concepts.

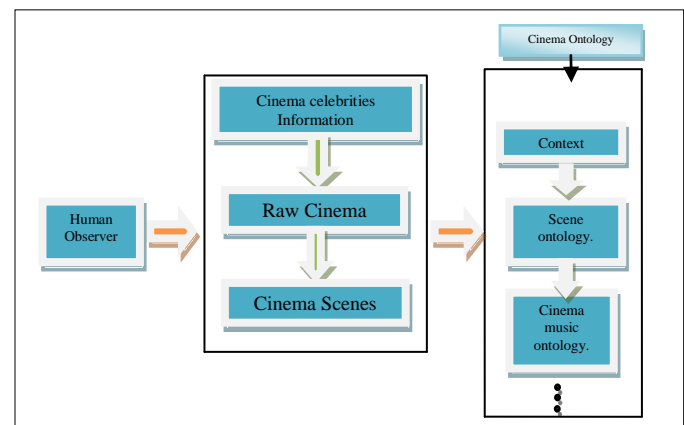


Figure.2 Construction of Cinema Ontology

The overall process can be formatted in the steps below.

Step1: Multiple cinemas are taken as the source of the semantic data. The power of ontology in organizing concepts can be used in modeling the video data.

Step2: Manual approach is adopted to identify the semantic concepts in cinemas. Cinemas are segmented into different scenes based on the concepts.

Step3: Identify the concept hierarchy. Identify the abstract, concrete concept classes in cinema scenes.

Step4: Concepts are classified into disjoint concepts, overlapping concepts, range concepts ...etc,.

Step5: Build the ontology using ontology construction tool - protégé and the ontology graph to visualize and evaluate the ontology.

Step6: Post-construction analysis by domain expert.

Multiple cinema are taken as a source, and based on which, core concepts, abstract concept class, concrete concept classes, concept instances and the concept hierarchy between them are identified. Manual annotation is generated for cinema scenes. The main aim is to extract the semantics of cinema scene using which semantic concepts can be detected and concept information can be maintained.

A. Identification of Concept in Cinema Domain

Concept represent Themes, Events, Actions, Locations, Emotion, Artist or anything that is desirable to mark the presence in the media object. Concepts may be organized in to hierarchies. The basic logical unit in the structure of cinema domain is Scene. Based on cinema, we have segmented the cinema into various Scene objects, which contains one complete meaningful scene. A cinema contains concepts like Themes, Events, Actions, Locations, Emotion, Artist ...etc,. A raw cinema V can be segmented in to n number of segments or video objects VO_i , i.e., $V = \{VO_1, VO_2, \dots, VO_i\}$ where i ranges from 1 to the number of scenes in the cinema. Each Cinema contains a number of concepts.

Let C be the set of all possible concepts in a given domain. $C = \{C_1, C_2 \dots C_j\}$ where j ranges from 1 to the possible number of concepts. The number of concepts and the type of concepts depends on the abstract view of the application and the user requirements. We now can view a raw video as a collection of concepts, $V = \{C_1, C_2 \dots C_i\}$. Each video object VO_i contains set of concepts C_c which is a sub set of the concept set C. $VO_i = \{C_1, C_6, C_y, C_j, \dots\}$. Concepts can be classified in to concept class based on the concept type. A concept can have z number of subclasses. For example, scene concept can be further classified into comedy, tragedy, fights, romance ...etc, based on the theme. Further a concept class can have number of concept values, $CC_m = \{CV_1, CV_2, \dots\}$, where CV_o is the possible values that the concept can have. For example action concept can have subclasses as Fight, comedy, Song, Tragedy ...etc,. Multimedia objects are described by a set of concepts $C_1, C_2, C_3, \dots, C_n$ where n is the number of concepts associated to cinema, each concept C_k can have m concept values. i.e., $VO_i = \{CC_1 (CV_1), CC_2 (CV_2), \dots, CC_n (CV_m)\}$. E.g.: $VO_i = \{\text{Song (romantic), Hero (Bachan), shot (trolley)}\}$. Concepts can be identified and added at any time which increases the flexibility of the proposed model. User can browse a cinema based on the semantic concepts like all comedy, tragedy, fighting, romantic ...etc and they can search specific type of comedy scene like comedy scenes in a song, comedy scenes in fighting ...etc,. [16][17][18] describes ontology tools. We have used Protégé as an Ontology developing tool to implement our cinema ontology construction [19] [20]. Protégé was developed by

(<http://protege.stanford.edu>) Mark Musen's group at Stanford University. We generated few figures with the onto graph plug-in to protégé as depicted in Fig.3.a, Fig.3.b, Fig.3.c. We selected OWL, as the ontology language, which is standard ontology language recommended by W3C.

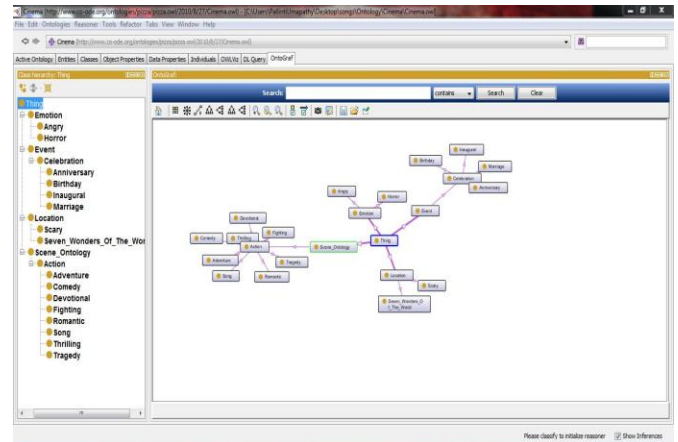


Figure 3.a Onto graph showing cinema ontology

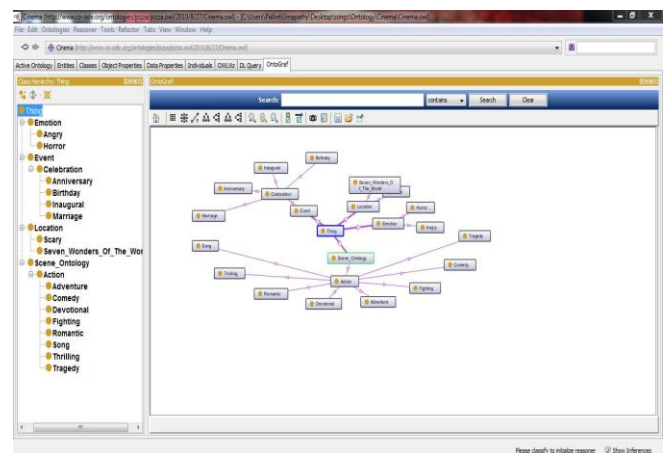


Figure 3.b Onto graph showing cinema ontology.

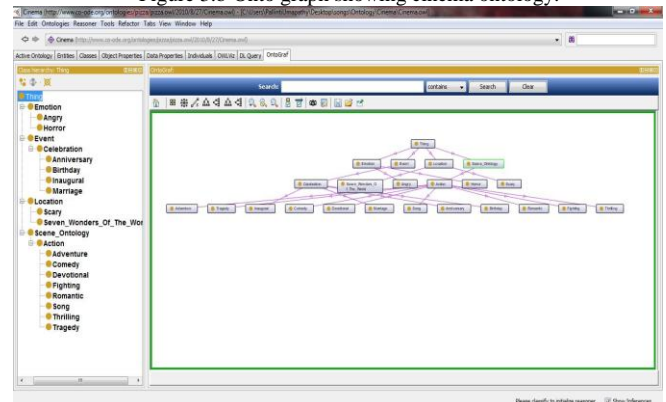


Figure 3.c Onto graph showing cinema ontology.

V. CONCLUSION AND FUTURE WORK

Ontology is widely accepted technique for knowledge system development. Ontology plays important role in design

and sharing of knowledge. Effective structure of the knowledge improves the efficiency of the retrieval system. Semantic web and ontology provide a method to construct and use resources by attaching semantic information to them. In this paper our focus is on construction of cinema ontology. Cinema ontology is defined by identifying the semantic concepts, context hierarchy and the ontology structure is presented. In the proposed approach based on the users and their requirement two sub ontology were developed. Cinema context ontology and cinema scene ontology. The former deals with the external information and while the later focus on the semantic concepts of the cinema scene and their hierarchy and the relationship among the concepts. Finally practical implementation of cinema ontology is illustrated using the protégé tool.

Further studies can be done towards:

- Designing and construction of information extraction system based on the cinema context ontology for extracting the context information and achieve the true scene of information sharing.
- Design and construction of ontology base cinema scene search engine which will support the cinema stake holder's needs by retrieving the appropriate cinema scenes pertaining to different themes, actors, actions etc.

The use of cinema ontology can more effectively support the construction of cinema scene library in television channels as well as cinema production companies for their cinema based programs and brings entertainment for cinema lovers.

ACKNOWLEDGMENT

This work has been partly done in the labs of Adhiyamaan College of Engineering where the first author is currently working as a Professor & Director in the department of Master of Computer applications. The authors would like to express their sincere thanks to Adhiyamaan College of Engineering for their support rendered during the implementation of this module.

REFERENCES

- [1] J. Chen, Q. M. Zhu, and Z. X. Gong, "Overview of Ontology-Based Information Extraction," *Computer Technology And Development*, vol. 17(10), pp. 84–91, 2007.
- [2] F. Gu, C. G. Cao, Y. F. Sui, and W. Tian, "Domain-Specific Ontology of Botany," *Journal of Computer Science and Technology*, vol. 19(2), pp. 238–248, 2004.
- [3] H. C. Jiang, D. L. Wang, and D. Z. Zhang, "Approach of Chinese Medicine Knowledge Acquisition Based on Domain Ontology," *Computer Engineering*, vol. 34(12), pp. 16–18, 21, 2008.
- [4] J. D. Yu, X. Y. Li, and X. Z. Fan, "Design and Implementation of Domain Ontology for Information Extraction," *Journal of University of Electronic Science And Technology of China*, vol. 37(5), pp. 746–749, 2008.

- [5] Gruber T.R: Towards principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, Volume 43, Issue 5-6, pp. 907-928, 1993.
- [6] Z. H. Deng, S. W. Tang, M. Zhang, D. Q. Yang, and J. Chen, "Overview of Ontology," *Acta Scientiarum Naturalium Universitatis Pekinensis*, vol. 38(5), pp. 229–231, 2002.
- [7] Zhou . M.Q , Geng G. H, and Huang S. G., "Ontology Development for Insect Morphology and Taxonomy System," *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pp. 324–330, December 2006.
- [8] N. F. Noy, and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05*, 2001.
- [9] Lili Zhao and Chunping Li, *Ontology Based Opinion Mining for Movie Reviews*, Third International conference, KSEM 2009, Vienna, Austria, November 2009 proceedings. Pp.204-214
- [10] Amancio Bouza [] : "MO – the Movie Ontology". MO – the Movie Ontology. 2010. movieontology.org
- [11] Shiyan Ou, Constantin Orasan, Dalila Mekhaldi, Laura Hasler, *Automatic Question Pattern Generation for Ontology-based Question Answering*.
- [12] Gijis Geleijnse A Case Study on Information Extraction from the Internet: Populating a Movie Ontology.
- [13] Sousan W.L, Wylie, chengxin, Chen, *Constructing Domain Ontology from Texts: A Practical Approach and a Case Study*, NWESP '09. Fifth International Conference on, 2009, Page(s): 98 – 101.
- [14] Xinli Zhao; Dongxia Zhao; Wenfei Gao; China Sci. & Technol. Exchange Center, Beijing, China Research on the construction of government ontology *Intelligent Computing and Intelligent Systems*, 2009. ICIS 2009. IEEE International Conference on 20-22 Nov. 2009 Pg 319 – 323.
- [15] Ling Zeng; Tonglin Zhu; Xin Ding; Study on Construction of University Course Ontology: Content, Method and Process *Computational Intelligence and Software Engineering*, 2009. CiSE , 2009 , Page(s): 1 - 4 .
- [16] F. L. Mariano and G. P. Asunción, "Overview and analysis of methodologies for building ontologies," *The Knowledge Engineering Review*, vol. 17(2), pp. 129–156, 2002.
- [17] Cimiano P., Volker J., and Studer R., "Ontologies on Demand? – A Description of the State-of-the-Art, Applications, Challenges and Trends for Ontology Learning from Text," *Information, Wissenschaft und Praxis*, Vol. 57, No. 6-7. (October 2006), pp. 315-320.
- [18] Duineveld A. et al. "Wonder Tools"? A Comparative Study of Ontological Engineering Tools." *Intl. Journal of Himian-Computer Studies*. Vol. 52 No. 6, pp. 11 11-1 133. 2000.
- [19] Michael Denny, *Ontology Building*, "A Survey of Editing Tools," <http://www.xml.com/pub/a/2002/11/06/ontologies.html>.
- [20] Matthew Horridge, Simon Jupp, Georgina Moulton, Alan Rector, Robert Stevens, Chris Wroe. *OWL Ontologies using protégé 4 and CO-ODE Tools Edition 1.1*. The University of Manchester , October 16, 2007.

AUTHORS PROFILE

Dr. Sunitha Abburu: Working as a Professor and Director, in the Department of Computer Applications, Adiyamaan College of Engineering, Tamilnadu, India. She received BSc and MCA from Osmania University, A.P, and India. M.phil and Ph.D from Sri Venkateswara University, A.P, India. She is having 13 years of teaching experience and 3 years of industrial experience.

Jinesh V N: (Graduate Member of Institution of Engineer's(India)) Obtained Diploma in Computer Science and Engineering from Board of technical studies, India; Bachelor of Engineering in Computer Science and engineering from The Institution of Engineer's (India) and M.Tech in Computer Science and Engineering from Visveswaraya Technological University, India. Currently he is working as a lecturer in Department of Computer science, The Oxford college of Science, Bangalore, India.

S-CAN: Spatial Content Addressable Network for Networked Virtual Environments

Amira Soliman, Walaa M. Sheta
Informatics Research Institute
Mubarak City for Scientific Research and Technology Applications
Alexandria, Egypt.

Abstract—Networked Virtual Environments (NVE) combines 3D graphics with networking to provide simulated environment for people across the globe. The availability of high speed networks and computer graphics hardware enables enormous number of users to connect and interact with each other. In NVE each node (user or avatar) should be aware of the existence and modification of all its neighbors. Therefore, neighborhood consistency that is defined as ratio between node's known and actual neighbors is a fundamental problem of NVE and should be attained as high as possible. In this paper, we address the neighborhood consistency by introducing S-CAN, a spatial Peer-to-Peer (P2P) overlay that dynamically organizes nodes in NVE to preserve spatial locality of users in NVE. Consequently, node's neighborhood will always maintain node's direct neighbors and hence node will be aware of other users and events within its visibility that is called node's Area-of-Interest.

Keywords: Networked Virtual Environments; Peer-to-Peer Systems; Interest Management.

I. INTRODUCTION

Networked virtual environment (NVE) [1, 2], also known as distributed virtual environment, is an emerging discipline that combines the fields of computer graphics and computer networks to allow many geographically distributed users interact simultaneously in a shared virtual environment. NVEs are synthetic worlds where each user assumes a virtual identity (called avatar) to interact with other human or computer players. Users may perform different actions such as moving to new locations, looking around at the surroundings, using items, or engaging in conversations and trades. Applications of NVE have evolved from military training simulation in the 80's to the massively multiplayer online games (MMOG) in the 90's [3, 4].

In NVEs each user is interested in only a portion of the virtual world called area-of-interest (AOI). All nodes in a node's AOI are said to be its neighbors. AOI is a fundamental NVE concept, as even though many users and events may exist in the system, each user, as in the real world, is only affected by nearby users or events. AOI thus specifies a scope for information which the system should provide to each user. It is thus essential to manage communications between users to permit receiving of relevant messages (generated by other users) within their AOI as they move around [5, 6].

As NVEs are shared environments, it is important that each participant perceives the same states and events. Consistency

in NVEs usually refers to consistent object states and event orderings [1], which are maintained through the transmission of event messages. In this paper, we focus on neighborhood consistency or topology consistency, which can be defined as the percentage of correctly known AOI neighbors. For example, a node that is aware of four out of five AOI neighbors, topology consistency is 80 percent [7]. In client/server NVE architectures, keeping high neighborhood consistency is trivial as all the user states are maintained by a centralized server. While, in P2P NVE, achieving neighborhood consistency is much harder as states are maintained by participating nodes [6].

Therefore, it is essential to dynamically organize P2P overlay network with respect to users' current positions in the virtual world by having each user connected to the geographically closest neighbors (users within AOI) [8]. In this paper we introduce the architecture of Spatial Content Addressable Network (S-CAN) for NVE. Our design is based on Content-Addressable Network (CAN) [9] for constructing P2P overlay.

CAN design centers around a virtual d -dimensional Cartesian coordinate space. CAN coordinate space is completely logical and has no relation to any physical coordinate system. However, in our P2P overlay we associate physical coordinate system with CAN coordinate space. So, physical location of users and objects in virtual environments determines their correspondent location in CAN coordinate space. The objective of this mapping relation between physical and CAN coordinates is to preserve the spatial locality among users and objects in NVE and hence attain user awareness.

The rest of this paper is organized as follows. Section 2 gives background overview of related work and CAN network overlay. Section 3 introduces the adaptations proposed in S-CAN. Experiments are presented in Section 4 with metrics and scenarios. Results are presented and discussed in Section 5. Conclusion and future work are given in Section 6.

II. BACKGROUND

A. Related Work

Various techniques have been proposed to address interest management in NVE. The earliest approached utilize multicast channels, where Virtual Environment (VE) is divided into regions and assign each region a multicast channel for notification messages propagation. Each avatar can subscribe to the channels of the regions overlapped with its AOI.

NPSNET [10], VELVET [11] and SimMUD [12] are examples of multicast NVE. However, this approach faces the inherent difficulty of determining the right region size. Since, too large regions deliver excessive messages to each avatar. While, small regions require many subscription requests and thus generate message overhead.

Approaches with spatial multicast messages have been developed to address the inadequacy of channel-based multicast. These approaches use Trees and Distributed Hash Tables (DHT) to store spatial relations among avatar and objects in NVE. Examples include N-trees [13], Solipsis [14] and VON [6]. However, these approaches maintain another data structure and protocol dedicated for interest management rather than the protocol used to develop the network overlay.

B. Content-Addressable Network (CAN)

CAN introduces a novel approach for creating a scalable indexing mechanism in P2P environments. It creates a logical d -dimensional cartesian coordinate space divided into zones, where zones are partitioned or merged as result of node joining and departure. The entire coordinate space is dynamically partitioned among all the nodes in the system such that every node "owns" its individual, distinct zone within the overall space. Fig. 1 shows a 2-dimensional $[0,1] \times [0,1]$ coordinate space partitioned between 5 nodes.

This virtual coordinate space is used to store (key, value) pairs as follows: to store a pair (K_1, V_1) , key K_1 is deterministically mapped onto a point P in the coordinate space using a uniform hash function. The corresponding (K_1, V_1) pair is then stored at the node that owns the zone within which the point P lies. To retrieve an entry corresponding to key K_1 , any node can apply the same deterministic hash function to map K_1 onto point P and then retrieve the corresponding value from the point P . If the point P is not owned by the requesting node or its immediate neighbors, the request must be routed through the CAN infrastructure until it reaches the node in whose zone P lies.

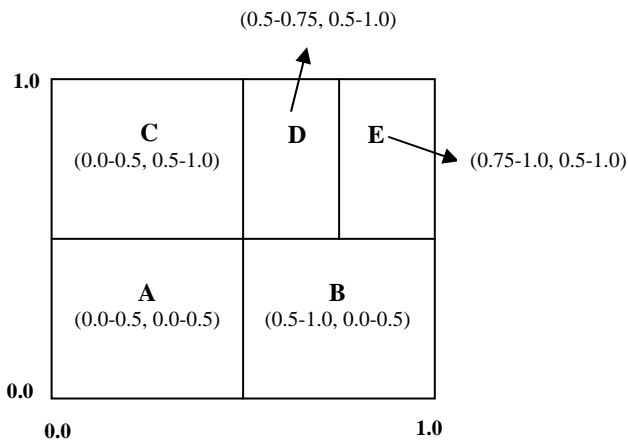


Figure 1. 2-d space with 5 nodes illustrating node's virtual coordinate zone.

1) Node Join: The entire space is divided amongst the nodes currently in the system. To allow the CAN to grow incrementally, a new node that joins the system must be

allocated its own portion of the coordinate space. This is done by an existing node splitting its allocated zone in half, retaining half and handing the other half to the new node. The process takes four steps:

1. The new node n must find a node already in S-CAN and send join request to it.
2. Next, using the S-CAN routing procedure, the join request is forwarded to the nearest node whose zone will be split.
3. Then, the neighbors of the split zone must be notified with new node.
4. Finally, move responsibility for all the keys and objects data files that are positioned in zone handed to n .

2) Routing: CAN nodes operate without global knowledge of the plane. Each node maintains a routing table consists of the IP addresses and logical zones areas of its immediate neighbors. In a d -dimensional coordinate space, two nodes are neighbors if their coordinate spans overlap along $d-1$ dimensions and abut along one dimension. For example, in fig. 1, node A is a neighbor of node B because its coordinate zone overlaps with A's along the Y axis and abuts along the X-axis. On the other hand, node D is not a neighbor of node A because their coordinate zones abut along both the X and Y axes.

Routing in CAN works by following the straight line path through the cartesian space from source to destination coordinates. Using its neighbor coordinate set, a node routes a message towards its destination by simple forwarding to the neighbor with coordinates closest to the destination coordinates. As shown in [9], the average routing path length is $(d/4)(n^{1/d})$ hops and individual nodes maintain $2 \times d$ neighbors. Thus, the number of nodes (and hence zones) in the network can grow without an increasing per-node-state, while the path length grows with $O(n^{1/d})$.

3) Node Departure: When nodes leave CAN, the zones they occupy must be taken over by the remaining nodes. The normal procedure for node leaving is to explicitly hand over its zone and associated (key, value) database to one of its neighbors. If the zone of one of the neighbors can be merging with the departing node's zone to produce a valid single zone, then this is done. The produced zone must have a regular shape that permits further splitting in two equal parts. If merge fails, the zone is handed to the neighbor whose current zone is the smallest, and that node will temporarily handle both zones.

III. THE PROPOSED OVERLAY S-CAN

As previously mentioned, our work leverages the design of CAN to support user awareness in NVEs. CAN constructs a pure logical coordinate plane to uniformly distribute data objects. In S-CAN, we use this characteristic to extend the logical plane with physical spatial meaning for distributing data objects based on spatial information. Furthermore, because user's AOI is often based on geographic proximity, we dynamically reorganize the network overlay with respect to

users' current positions in the VE. Therefore, as users move inside VE, the coordinates of their assigned zones and neighborhood map will be changed to reflect their current positions. In section (A) we illustrate the overlay adaptation process. Next, in section (B) we present the stabilization procedure.

A. Overlay Construction

S-CAN divides the coordinate space into zones according to number of nodes and their locations. Moreover, each node maintains a routing table (called neighbor map) that stores adjacent neighbors with their associated zones. Node's avatar moves freely as long as it is still located within node associated zone. In case of any movement outside zone coordinates, node has to change its location in network overlay (which means that node moves within network overlay). This node movement will be done by performing node departure then rejoin according to avatar new location. Therefore, when node changes its location, its associated zone will be changed and new AOI neighbors will be stored in its routing table.

Each node maintains a neighbor map of the following data structure:

*HashMap {Direction ,
HashMap {NodeID , ZoneBoundary [] []}}*

Direction takes a single value from {"East", "West", "North", "South"}. Where, ZoneBoundary is a 2-d array storing values of start and end in x and y direction respectively.

In our proposed overlay, we differentiate between two types of neighborhood based on number of neighbors sharing the same border line. In the first type, there is only one neighbor sharing all border line, where, in type two there are more than one neighbor. In order to determine neighborhood type, there are two methods developed to calculate overlap direction between zones. The first method is *overlapAllBoundary* that returns direction where zones share from start to end for example as in fig. 1, calling *overlapAllBoundary* between zones of nodes A and B returns "East" as node B is in east direction of node A. If we call the method with nodes reversed (that is B then A), it will return direction "West". While, the second method *overlapPartBoundary* returns the direction where two zone share a part from it. In fig. 1, there is *overlapPartBoundary* between nodes B and D in "North" direction.

After avatar's movement, node performs algorithm (1) mentioned below. First, it compares avatar new position with the boundary of the current associated zone. If the new position is out of zone boundary (line 3), it searches within its neighbors to find a neighbor with a valid zone to merge (lines 4:11). We use *overlapAllBoundary* method (line 5) in order to verify generating a regular zone shape at the end of merge process. When finding a matched neighbor, a merge request will be sent to it and node freezes till a response is received. The received merge response indicates whether merge process succeeds or fails. Neighbor rejects merge request as it is currently executing a process that is going to change its associated zone like splitting for new joined node, or in movement processing and trying to hand over its zone.

Algorithm1. Avatar movement process(posX, posY)

```
1: /*psoX is the new user's X coordinate*/
2: /*psoY is the new user's Y coordinate*/
3: if not inMyZone(posX, posY) then
4:   for each neighbor N ∈ myNeighbors do
5:     if overlapAllBoundary(zone, N.zone) do
6:       sendMerge(zone, N)
7:       if mergeSucceed() do
8:         break
9:     end if
10:   end if
11: end for
12: if not mergeSucceed() do
13:   sendAddUnallocatedZone(zone)
14: end if
15: join(posX, posY)
16: end if
```

Subsequently, after merge took place, the node sends a join request to one of its oldest neighbors existing in its move direction (line 15). Then, a join request will be forwarded till reaching the node that will accept and split its zone with the requesting node. Furthermore, the node that performs merge will be responsible for notifying overlapped neighbors with the change happened. So that, it will forward two messages to overlapped neighbors, the first indicates its new zone coordinates, while the second message notifies neighbors to delete the departed node.

However, not all merge requests succeed, so in the next section we illustrate the process performed in case of merge failure and coordinate stabilization process.

B. Stabilization

When nodes move in S-CAN, we need to ensure that their associated zones are taken by the remaining nodes. Nevertheless, the merge process succeeds only when the zone of one of the neighbors can be merged with the moved node's zone to produce a valid single zone. If not, the zone is declared as an unallocated zone and handed to a specific node that is responsible for managing unallocated zones. This node is known as Rendezvous node. Rendezvous node serves as a bootstrap node in its region. It is a static node and is launched with the system start. This node maintains two lists, one for storing the unallocated zones and the other for listing the avatars located in its region.

When a new unallocated zone is received, Rendezvous node verifies if this zone can be merged with one of the old zones in order to minimize scattering in the coordinate space. It iterates on the existing zones and uses the *overlapAllBoundary* function to check if there is any overlap as shown in algorithm (2). If merge can be made, it removes the old zone and adds the new merged one to the unallocatedZones list (lines 3:10). Otherwise, add the zone will be added to the list of unallocatedZones (lines 11:13).

When Rendezvous node receives a join request, it first verifies if requesting node position lies in the coordinates of one of the unallocated zones. If so, it sends the reply with the unallocated zone coordinates. Then, in order to let the new joined node

know its neighbors, the Rendezvous node performs a neighbor search over its avatars list as illustrated in algorithm (3).

Algorithm2. Unallocated zones merging(pZone)

```
1: /* pZone is the new unallocated zone found*/
2: mergeDone = false
3: for each unallocated zone Z ∈ unallocatedZones do
4:   if overlapAllBoundary(pZone, Z) then
5:     newZ = merge(pZone, Z)
6:     mergeDone = true
7:     unallocatedZones.remove(Z)
8:     unallocatedZones.add(newZ)
9:   end if
10: end for
11: if not mergeDone then
12:   unallocatedZones.add(pZone)
13: end if
```

Algorithm3. Search for neighbors(zone, n)

```
1: /* zone is the unallocated zone associated to node n*/
2: /* n is the profile information of node n*/
3: for each mobile node M ∈ mobileNodes do
4:   if overlapPartBoundary(zone, M.zone) then
5:     sendAddNeighbor(n, M)
6:     sendAddNeighbor(M, n)
7:   end if
8: end for
```

IV. EXPERIMENTAL EVALUATION

A. Experimental Setup

We build S-CAN prototype using JXTA framework [15]. JXTA is open-source project that defines a set of standard protocols for ad-hoc P2P computing. JXTA offers a protocol suite for developing a wide variety of decentralized network application [16]. We generate experimental data set using JXTA IDFactory generator, this data set includes nodes' ID and objects' ID. Then, those IDs are mapped to coordinate space to generate nodes and objects physical locations in coordinate space.

In each experiment the nodes start by loading S-CAN service and then join the network overlay. Initially we divide the coordinate space into four regions. Each region contains a Rendezvous node that is used as bootstrap in this region. For any node to join overlay, it sends join request to Rendezvous node in its region. Then, the request is forwarded till reaching the nearest node whose zone will be split. Fig. 2 shows the coordinate space after the joining of 40 nodes. Green dots stands for object data files, and blue dots stands for avatars, while, the red dots stands for Rendezvous nodes. After node joining, avatars start to move. We move the avatars using two navigation patterns: random and spiral patterns as shown in fig. 3.

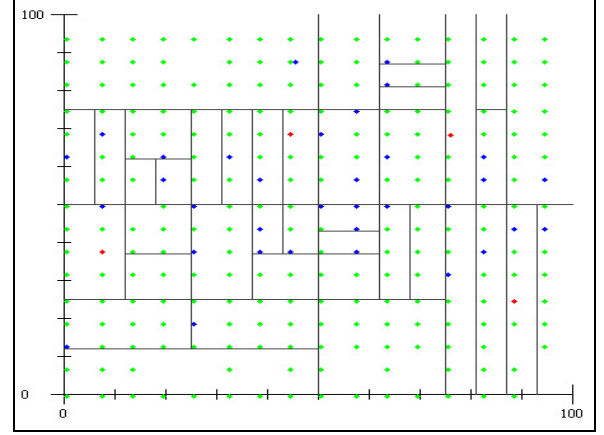


Figure 2. 2-d space with 40 nodes illustrating zone splitting based on avatars' locations.

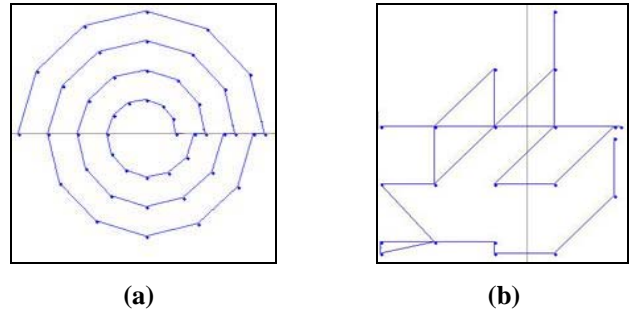


Figure 3. Avatars navigation patterns: (a) spiral pattern, (b) random pattern.

B. Performance Metrics

In order to evaluate the performance of the proposed prototype, we use the following factors:

Number of hops: presents the number of nodes in the routing path of a request message. It presents the number of nodes contacted on the way from source to destination. We measure number of hops for join and get requests.

Number of files transferred: indicates the number of files transmitted to nodes after join or move process. Those files are objects' data files associated with node's zone or scene files need to be loaded by avatar.

Number of update messages: presents the number of messages sent after node's move to reorganize network overlay. We count messages sent to reflect zone coordinates changes and neighborhood map updates.

Number of unallocated zones: describes the number of zones exist in unallocated zones lists. Furthermore, we count number of message sent to search for neighbors after re-assigning of unallocated zones.

AOI notification: is defined as the number of hops taken to notify all the neighbors in node's AOI with the change occurred.

V. RESULTS

A. Number of Hops

Number of hops reflects the request delay caused by the underlying P2P overlay routing. We performed different experiments with different number of nodes (20, 40, and 80 nodes). After node joining and receiving files associated with its zone, it starts to load files of data objects located in its current scene. If there is a missing file needed to be loaded, it forwards a get request to one of its neighbors according to missing object location. Table 1 shows the average number of hops obtained per each network overlay size for join and get queries.

Table 1. Number of hops per join and get requests with different number of nodes

	No. of hops (join)	No. of hops (get)
20 Nodes	2	1
40 Nodes	3	3
80 Nodes	4	4

The results indicate that join complexity in S-CAN is lower than join complexity in CAN. As in CAN routing complexity is $O(n^{1/d})$ which gives 5, 7, and 9 for 20, 40, and 80 nodes respectively. The reason behind this is in S-CAN we have four bootstrap nodes (Rendezvous nodes) and node send join request to bootstrap node in its region which is somehow near to it.

It is also apparently that as the number of nodes in system increases, as the size of associated zones decreases. Hence, number of hops of get request increases with increase of number of nodes. Table 1 shows that in 20 nodes, any get request can be served from direct neighbor (that is single hop message).

B. Number of Files Transmitted

This factor illustrates how the number of objects in VE affects zone merging and splitting in NVE. Since, with each node movement, node sends files located in previous zone and receives files associated with new zone.

Fig. 4 shows the number of received objects after join with different scene sizes (in terms of number of objects in VE) and nodes in NVE. Moreover, fig. 5 explores number of missed scene objects with different scene sizes and nodes. It is clear that as the number of nodes increases (smaller associated zone) as the number of received files decreases. However, as the size of associated zone decreases as the number of missed scene objects increases (as shown in fig. 5). So, nodes have to send more get requests to get missed objects from neighbors. So, we can conclude that there is a trade-off between zone size and scene size.

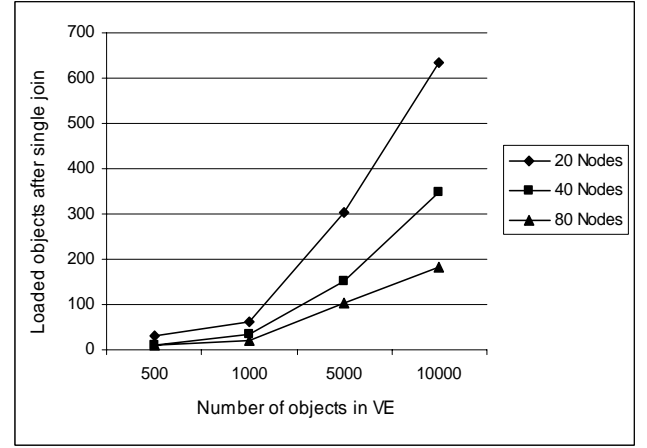


Figure 4. Number of data files received with different number of nodes and objects.

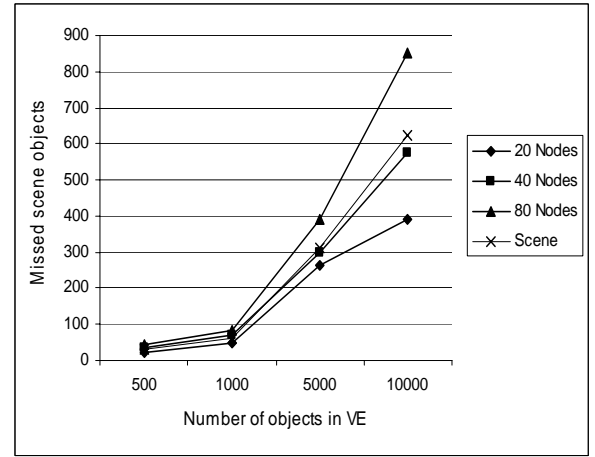


Figure 5. Number of missed scene objects with different number of objects and nodes.

C. Number of Update Messages

In order to reorganize overlay network with nodes movements, a number of messages are sent to notify existing nodes with current updates. Those updates are classified into two different categories: zone coordinates updates and neighbor map updates. First category covers changes occurred after merging old zone of moving node and splitting after rejoining based on new location. While second category covers changes in neighbor map for removing the moved node and adding it after rejoin. Based on accepted merge response (as illustrated in algorithm 1), the moving node sends a remove neighbor request to its neighbor to delete it and add neighbor that accepts merge with the new zone coordinate. The neighbor that accepts merge will be responsible for notifying its neighbors with new zone coordinates. Finally, after node rejoining, the neighbors of the split zone must be notified with new node.

Fig. 6 explores the average number of messages sent to reflect a single node movement with different number of nodes. Update zone indicates number of messages sent to update zone coordinates, while Update neighbor map indicates number of messages sent to update neighbor map. The last column

indicates the routing complexity in CAN. We add it to the figure to illustrate that node movement complexity can be considered as the same as routing complexity in CAN.

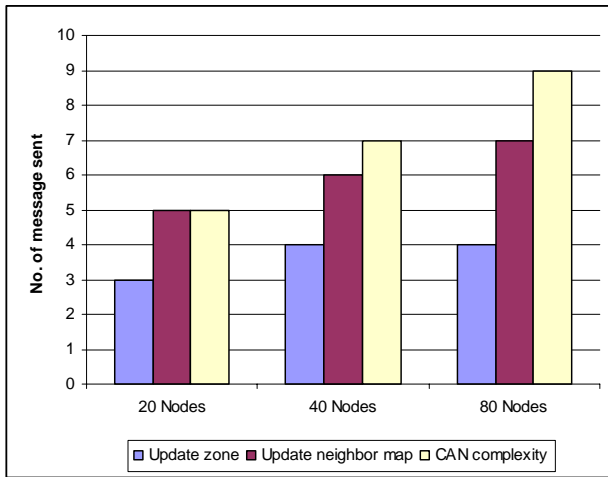


Figure 6. Number of messages sent to reflect a single node movement.

D. Number of Unallocated Zones

In this experiment, we count total number of nodes movement, resulted unallocated zones, reassigned unallocated zones, and finally total number of messages sent to fix neighborhood after reassigning unallocated zones. Table 2 lists the results obtained.

Table 2. Number of unallocated zones found and neighbor search queries sent.

	No. of moves	Unallocated zones	Re-assigned	Neighbor search
20 Nodes	112	33	13	66
40 Nodes	252	77	22	98
80 Nodes	532	161	37	149

From the result obtained, we can figure out that the rate of adding and reassigning an unallocated zone is almost the same with different number of nodes in the overlay. Therefore, we can conclude that there is no relation between zone size and growth of unallocated zones in coordinate space. Moreover, the average number of neighbor search queries per single reassignment of an unallocated zone is lower than the routing complexity of CAN.

E. AOI Notification

The objective of this experiment is to study the number of hops that event's message takes till reaching all the neighbors in node's AOI. When avatar changes a property of any objects in its scene, node calculates AOI boundary of this objects and sends a notification message to neighbors whose zone overlaps with that AOI boundary. Upon receiving this message, the receiving node on its turn will forward it to its neighbors whose zones overlap with AOI boundary. Therefore, message will be forwarded till reaching all neighbors within the first node's AOI.

Fig. 7 explores number of hops that event message takes with different number of nodes in overlay and different values of AOI radius. The obtained results show that zone size affects

number of hops taken by event message. Since, as large as the zone size is, as fast as message reaches all neighbors in node's AOI.

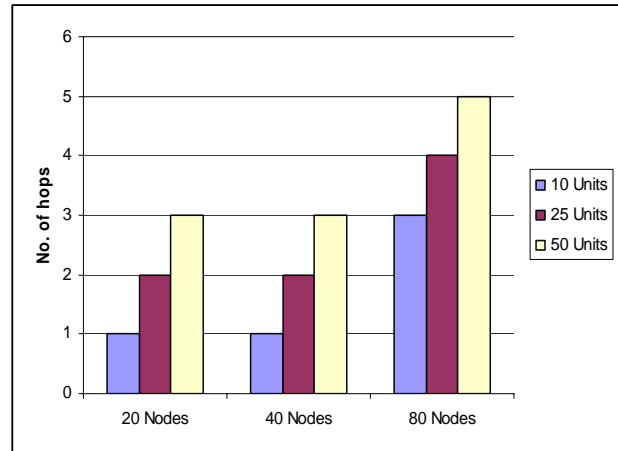


Figure 7. Number of hops taken to notify all neighbors in node's AOI with different AOI radius.

VI. CONCLUSION AND FUTURE WORK

P2P systems have generated intense interest in the research community because their robustness, efficiency, and scalability are desirable for large-scale systems. We have presented S-CAN, a spatial P2P network overlay that preserves both spatial locality and neighborhood consistency in NVEs. We have presented S-CAN system operations namely overlay construction and stabilization. We perform set of experiments to measure S-CAN performance against set of common factors such as number of hops, number of files transferred, and AOI notification. Results show that we can achieve both AOI and neighborhood consistency.

We plan to extend our work in several directions. First, we will investigate the effect of node failure and message loss on network overlay construction and stabilization. Since, missing updates will lead to some nodes do not know all of their neighbors and in this situation NVE will work abnormally. Second, we will study using start and stop levels of zone splitting in coordinate to minimize the cost node movements. We expect that limiting zone splitting to a specific size and assign new node a mirror of zone rather than splitting will enhance the overall performance as it will minimize message sent to update nodes' neighbor map.

ACKNOWLEDGEMENT

This project is funded by the Egyptian Ministry of Communication and Information Technology under grant "Development of virtual Luxor" project.

REFERENCES

- [1] S. Singhal and M. Zyda. Networked Virtual Environments: Design and Implementation. ACM Press/Addison-Wesley Publishing, 1999.
- [2] J. Smed, T. Kaukoranta, and H. Hakonen, "Aspects of Networking in Multiplayer Computer Games," in Proc. ADCOG. Nov. 2001, pp. 74-81.
- [3] D. C. Miller and J. A. Thorpe, "SIMNET: The Advent of Simulator Networking," Proc. IEEE, vol. 83, no 8, pp. 1114-1123, Aug. 1995.

- [4] T. Alexander, "Massively Multiplayer Game Development," Charles River Media, 2003.
- [5] S.-Y. Hu, J.-F. Chen, and T.-H. Chen, "Von: A scalable peer-to-peer network for virtual environments," *IEEE Network*, vol. 20, no. 4, 2006.
- [6] J. Jiang, J. Chiou, and S. Hu, "Enhancing Neighborhood Consistency for Peer-to-Peer Distributed Virtual Environments," in *Proc. of the 27th international Conference on Distributed Computing Systems Workshops*, Jun 22 - 29, 2007.
- [7] Y. Kawahara, T. Aoyama, and H. Morikawa, "A Peer-to-Peer Message Exchange Scheme for Large-Scale Networked Virtual Environments," *Telecomm. Sys.*, vol. 25, no. 3-4, 2004, pp. 353-70.
- [8] R. Cavagna, M. Abdallah, and C. Bouville, "A framework for scalable virtual worlds using spatially organized P2P networks," in *Proc. of the 2008 ACM Symposium on Virtual Reality Software and Technology*, Oct 27 - 29, 2008.
- [9] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, "A Scalable Content-Addressable Network," *ACM SIGCOMM Conference*, 2001.
- [10] M. R. Macedonia, M. J. Zyda, D. R. Pratt, D. P. Brutzman, and P. T. Barham, "Exploiting reality with multicast groups," *IEEE Computer Graphics and Applications*, vol. 15, no. 5, pp. 38-45, 1995.
- [11] J. C. Oliveira and N. D. Georganas, "Velvet: An adaptive hybrid architecture for very large virtual environments," *Presence*, vol. 12, no. 6, pp. 555-580, 2003.
- [12] B. Knutsson, H. Lu, W. Xu, and B. Hopkins, "Peer-to-peer support for massively multiplayer games," *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 1, pp. -107, Mar 2004.
- [13] C. Gauthier Dickey, V. Lo, and D. Zappala, "Using n-trees for scalable event ordering in peer-to-peer games," in *Proc. of the international Workshop on Network and Operating Systems Support For Digital Audio and Video*, Jun 13 - 14, 2005.
- [14] J. Keller and G. Simon, "Solipsis: A massively multi-participant virtual world," in *PDPTA*, 2003.
- [15] JXTA Home Page, <http://www.jxta.org>
- [16] S. Oaks, B. Traversat, and L. Gong, "JXTA in a Nutshell," O'Reilly Press, 2002.

AUTHORS PROFILE

Wala M. Sheta is an associate professor of Computer graphics in Informatics Research Institute at Mubarak city for Scientific Research (MUCSAT) since 2006. During 2001-2006 he has worked as Assistant professor at MUCSAT. He holds a visiting professor position at University of Louisville in US and University of Salford in UK. He advised approximately 20 master's and doctoral graduates, his research contributions and consulting spans the areas of real-time computer graphics, Human computer Interaction, Distributed Virtual Environment and 3D image processing. He participated and led many national and multinational research funded projects. He received M.Sc. and PhD in Information Technology from University of Alexandria, in 1993 and 2000, respectively. He received B.Sc. from Faculty of Science, University of Alexandria in 1989.

Amira Soliman is an assistant researcher at Informatics Research Institute at MUCSAT. She Received the M.Sc. in computer science from Faculty of Computers, Cairo University in 2010. Amira's research interests include P2P Systems, Multi-Agent Systems, Software Engineering, Semantic and Knowledge Grids, Parallel Computing, and Mobile Applications.

Combinatory CPU Scheduling Algorithm

Saeeda Bibi¹, Farooque Azam¹, Yasir Chaudhry²

¹Department of Computer Engineering

College of Electrical and Mechanical Engineering,
National University of Science and Technology, Islamabad, Pakistan

²Department of Computer Science

Maharishi University of Management
Fairfield, Iowa USA

Abstract—Central Processing Unit (CPU) plays a significant role in computer system by transferring its control among different processes. As CPU is a central component, hence it must be used efficiently. Operating system performs an essential task that is known as CPU scheduling for efficient utilization of CPU. CPU scheduling has strong effect on resource utilization as well as overall performance of the system. In this paper, a new CPU scheduling algorithm called Combinatory is proposed that combines the functions of some basic scheduling algorithms. The suggested algorithm was evaluated on some CPU scheduling objectives and it was observed that this algorithm gave good performance as compared to the other existing CPU scheduling algorithms.

Keywords—component: Operating System, CPU scheduling, First Come First Serve Algorithm, Shortest Job First Algorithm,

I. INTRODUCTION

Operating system performs variety of tasks in which scheduling is one of the basic task. All the resources of computer are scheduled before use; as CPU is one of the major computer resources therefore its scheduling is vital for operating system [1]. When more than one process is ready to take control of CPU, the operating system must decide which process will take control of CPU first. The component of the operating system that is responsible for making this decision is called scheduler and the algorithm used by it is called scheduling algorithm [2].

In computer system, all processes execute by alternating their states between two burst cycles; CPU burst cycle and I/O burst cycle. Generally, a process starts its execution with a CPU burst then performs I/O (I/O burst), again another CPU burst then another I/O burst and this alternation of burst cycle continues until the completion of the process execution. CPU bound process is that which performs a lot of computational tasks and do little I/O while I/O bound process is that which performs a lot of I/O operations [1]. The typical task performed by the scheduler is to give the control of CPU to another process when one process is doing the I/O operations. The

reason behind it is that I/O takes long time to complete its operation and CPU has to remain idle [3, 4].

There are three different types of schedulers that are working in the operating system. Each scheduler has its own tasks that differentiate it from the others. These are:

A. Long-term Scheduler

It is also called high level scheduler, admission scheduler or job scheduler. It works with the job queue or high level queue and decides which process or job to be admitted to the ready queue for execution. Thus, the admission of the processes to the ready queue for execution is controlled by the long-term scheduler [5]. The major objective of this scheduler is to give balanced mix of jobs i.e. CPU bound and I/O bound, to the short-term scheduler [6].

B. Medium-term Scheduler

It is also called mid-term scheduler. This scheduler is responsible to remove the processes from main memory and put them in the secondary memory and vice versa. Thus, it decreases degree of multiprogramming. This is usually known as swapping of processes (“swapping-in” or “swapping out”) [5].

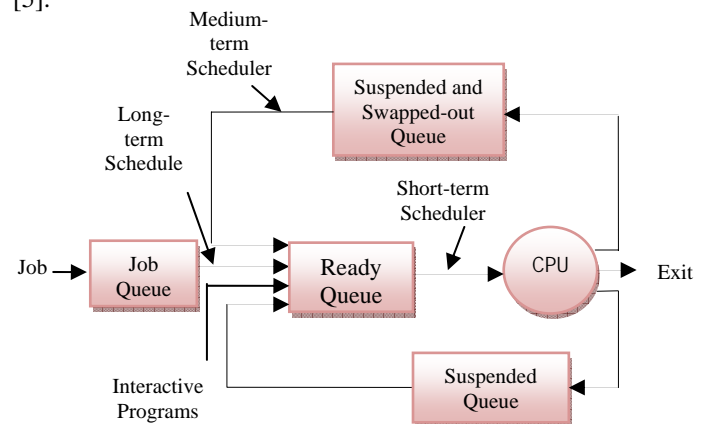


Figure 1: Schedulers

C. Short-term Scheduler

It is also called dispatcher or CPU scheduler. It decides which process from the ready queue takes control of the CPU next for execution [1]. Short-term scheduler makes scheduling decision much more frequently as compared to the other two schedulers. This decision is made on the basis of two disciplines these are non-preemptive and preemptive. In non-preemptive, the scheduler is unable to take control of the CPU forcefully from the processes. Processes take control of the CPU until the completion of execution. In preemptive, the scheduler is able to take control of the CPU forcefully from the processes when it decides to take CPU to the other process [5].

Design of CPU scheduling algorithm affects the success of CPU scheduler. CPU scheduling algorithms mainly depends on the criteria; CPU utilization, throughput, waiting time, turnaround time and response time [5]. Consequently, the major attempt of this work is to develop an optimal CPU scheduling algorithm that is suited for all types of processes and gives fair execution time to each process.

The organization of rest of the paper is as follow: Section II discusses existing scheduling algorithms. Section III describes proposed scheduling algorithm. Section IV contains pseudo code of the algorithm. Experimental Evaluation & Results have been given in Section V followed by conclusion.

II. OVERVIEW OF EXISTING CPU SCHEDULING ALGORITHMS

The basic CPU scheduling algorithms, their advantages and disadvantages are discussed in this section.

A. First Come First Served (FCFS) Scheduling

It is the simplest CPU scheduling algorithm that permits the execution of the process on the basis of their arrival time means the process having earlier arrival time will be executed first. Once the control of CPU is assigned to the process, it will not leave the CPU until it completes its execution. For small processes this technique is fair but for long processes it is quite unfair [7].

This algorithm is simple and can be implemented easily using FIFO queue. The problems of this algorithm are: the average waiting time, average turnaround time and average response time are high therefore it is not suitable for real time applications [9]. A long burst time process can monopolize CPU, even if burst time of other process is too short called convoy effect. Hence throughput is low [8].

B. Shortest Job First (SJF) Scheduling

This algorithm is non-preemptive in nature and permits the processes to execute first that have smaller burst time [10]. If more than one process has same burst time then control of CPU is assigned to them on the basis of First Come First Served. In most system, this algorithm is implemented for maximum throughput [5].

SJF algorithm is an optimal scheduling algorithm; it gives minimum average waiting time and average turnaround time [11] because it executes small processes before large ones. The difficulty of this algorithm is to know length of CPU burst of next process and it is usually unpredictable [9], there is also a problem of starvation in this algorithm because the arrival of processes having short CPU burst prevents processes having long CPU burst to execute [5].

C. Round Robin (RR) Scheduling

In this algorithm, a small unit of time called time quantum or time slice is assigned to each process. According to that time quantum processes are executed and if time quantum of any process expires before its complete execution, it is put at the end of the ready queue and control of the CPU is assigned to the next incoming process.

Performance of Round Robin totally depends on the size of time quantum. If size of time quantum is too small; it will cause many context switches and also affect the CPU efficiency. If time quantum is too large; it will give poor response time that approximately equal to FCFS [1]. This algorithm is preemptive in nature [7] and is suitable for time sharing systems. Round Robin algorithm gives high waiting time therefore deadlines are rarely met in it [5].

D. Priority Based Scheduling

In this algorithm, priority is associated with each process and on the basis of that priority CPU is allocated to the processes. Higher priority processes are executed first and lower priority processes are executed at end [4]. If multiple processes having the same priorities are ready to execute, control of CPU is assigned to these processes on the basis of FCFS [1, 3].

In this algorithm, average waiting time and response time of higher priority processes is small while waiting time increases for processes having equal priority [5, 12]. The major problem with this algorithm is problem of starvation that can be solved by a technique called aging [1].

E. SJRR CPU Scheduling Algorithm

In this algorithm, all the incoming processes are sorted in ascending order in the ready queue. Time quantum is

calculated and assigned to each process. On the basis of that time quantum, processes are executed one after another. If time quantum expires, CPU is taken from the processes forcefully and assigned to the next process; the preempted processes are put at the end of the ready queue [7].

SJRR provides fair share to each process and is useful in time sharing systems. It provides minimum average time and average turnaround time [7]. The problem with this algorithm is that if calculated time quantum is too small then there is overhead of more context switches.

III. PROPOSED SCHEDULING ALGORITHM

In this algorithm, a new factor F is calculated that is addition of two basic factors (arrival time and burst time of the processes). Here is the equation that shows this relation:

$$F = \text{Arrival Time} + \text{Burst Time}$$

This factor F is assigned to each process and on the basis of this factor processes are arranged in ascending order in the ready queue. Processes having highest value of the factor are executed first and those with lowest value of the factor are executed next. Depend on this new factor CPU executes the process that:

- Has shortest burst time
- Submit to the system at start

Proposed CPU scheduling algorithm reduces waiting time, turnaround time and response time and also increases CPU utilization and throughput. It has resolved the problem of starvation at much more extent and there is no problem of context switching in this algorithm.

The working of the proposed algorithm is as given below:

1. Take list of processes, their burst time and arrival time.
2. Find the factor F by adding arrival time and burst time of processes.
3. On the basis of factor, arrange processes and their relative burst time in ascending order using any sorting technique.
4. Calculate waiting time of each process.
5. Iterate through the list of processes
 - a. Add total waiting time with waiting time of each process to find total waiting time
 - b. Add burst time and waiting time of each process to find turnaround time
 - c. Add total turnaround time and turnaround time of each process to find total turnaround time
6. Average waiting time is calculated by dividing total waiting time with total number of processes.
7. Average turnaround time is calculated by dividing total turnaround time with total number of processes.

IV. PSEUDO CODE

```
f ← 0
temp ← 0
total_tatime ← 0.0
tw_time ← 0.0
avg_wt ← 0.0
avg_tatime ← 0.0

For i ← 0 to process
    F[i] ← atime[i] + btime[i]

For i ← process-1 to 0
    For j ← 1 to process
        IF F[j-1] > F[j]
            f ← F[j-1]
            F[j-1] ← F[j]
            F[j] ← f
            temp ← btime[j-1]
            btime[j-1] ← btime[j]
            btime[j] ← temp
            ptemp ← proname[j-1]
            proname[j-1] ← proname[j]
            proname[j] ← ptemp

wtime[1] ← 0
For j ← 1 to count
    wtime[j] ← btime[j-1] + wtime[j-1]

For j ← 0 to process
    tw_time ← tw_time + wtime[j]
    tatime[j] ← b[j] + wtime[j]
    total_tatime ← total_tatime + tatime[j]

avg_wt ← tw_time / process
avg_tatime ← total_tatime / process
```

V. EXPERIMENTAL EVALUATION & RESULTS

To explain the performance of proposed scheduling algorithm and to compare its performance with the performance of existing algorithms; consider the following set of processes along with their burst time, arrival time in milliseconds and priority in numbers as shown in the Table 1:

Process Name	Arrival Time	Burst Time	Priority
P1	0	20	6
P2	1	10	8
P3	2	3	2
P4	3	13	1
P5	4	10	4

Table 1: Set of Processes

Proposed CPU scheduling algorithm was implemented with existing CPU scheduling algorithms and performed detailed analysis by using Deterministic Evaluation method. Following Gantt charts of each algorithms, average waiting time and average turnaround time was obtained from this method.

A. Gantt Chart:

a. First Come First Served Scheduling:

P1	P2	P3	P4	P5	
0	10	12	21	24	29

Figure 2: Gantt chart for FCFS

b. Shortest Job First Scheduling:

P2	P4	P5	P3	P1	
0	2	5	10	19	29

Figure 3: Gantt chart for SJF

c. Round Robin Scheduling:

Here time quantum assigns to each process is 8.

P1	P2	P3	P4	P5	P1	P3	
0	8	10	18	21	26	28	29

Figure 4: Gantt chart for RR

d. Priority Based Scheduling:

P3	P2	P1	P4	P5	
0	9	11	21	24	29

Figure 5: Gantt chart for Priority Scheduling

e. SJRR Scheduling:

P2	P4	P5	P3	P1	P3	P1	
0	2	5	10	15	20	24	29

Figure 6: Gantt chart for SJRR Scheduling

f. Proposed Combinatory CPU Scheduling:

P2	P4	P5	P1	P3	
0	2	5	10	20	29

Figure 7: Gantt chart for Proposed Combinatory Scheduling

B. Waiting Time:

Waiting Time of the process is calculated as the time taken by the process to wait for the CPU in the ready queue. From the Gantt chart of the proposed Combinatory Scheduling, it is observed that waiting time for the processes P2, P4, P5, P1 & P3 is 0, 2, 5, 10 & 20 respectively and the average waiting time is $(0+2+5+10+20) / 5 = 7.4\text{ms}$. Waiting Time for all other algorithms is calculated in the same way. Table 2 shows

waiting time of each process and average waiting time for each scheduling algorithm.

C. Turnaround Time:

Turnaround Time of the process is calculated as the interval between the time of the submission of the process to the time of the completion of that process. From the Gantt chart of the proposed Combinatory Scheduling, it is observed that turnaround time for the processes P1, P2, P3, P4 & P5 is 20, 2, 29, 5 & 10 respectively and the average turnaround time is $(20+2+29+5+10) / 5 = 13.2\text{ms}$. Turnaround Time for all other algorithms is calculated in the same way. Table 3 shows turnaround time of each process and average turnaround time for each scheduling algorithm.

Process Name	Waiting Time (ms)					
	FCFS	SJF	RR	Priority	SJRR	Proposed Algorithm
P1	0	19	18	11	19	10
P2	10	0	8	9	0	0
P3	12	10	20	0	15	20
P4	21	2	18	21	2	2
P5	24	5	21	24	5	5
Avg. Waiting Time	13.4	7.2	17	13	8.2	7.4

Table 2: Waiting Time of each process and Average Waiting Time for Each Scheduling Algorithm

Process Name	Turnaround Time (ms)					
	FCFS	SJF	RR	Priority	SJRR	Proposed Algorithm
P1	10	29	28	21	29	20
P2	12	2	10	11	2	2
P3	21	19	29	9	24	29
P4	24	5	21	24	5	5
P5	29	10	26	29	10	10
Avg. Turnaround Time	19.2	13	22.8	18.8	14	13.2

Table 3: Turnaround Time of each process and Average Turnaround Time for Each Scheduling Algorithm

VI. CONCLUSION

The proposed algorithm along with existing algorithms has been simulated with C#.NET code and comparisons are made between the performance of proposed algorithm and existing algorithms. Graphical representation of these comparisons is shown in Figure 8 and Figure 9.

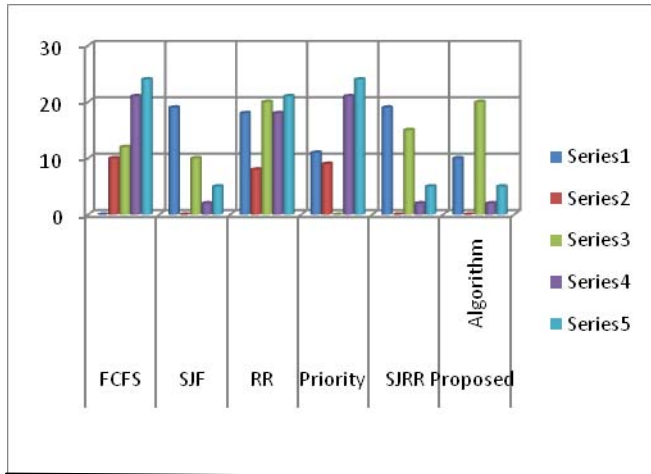


Figure 8: Comparison of Waiting Time of Proposed Algorithm with Waiting Time of Existing Algorithms

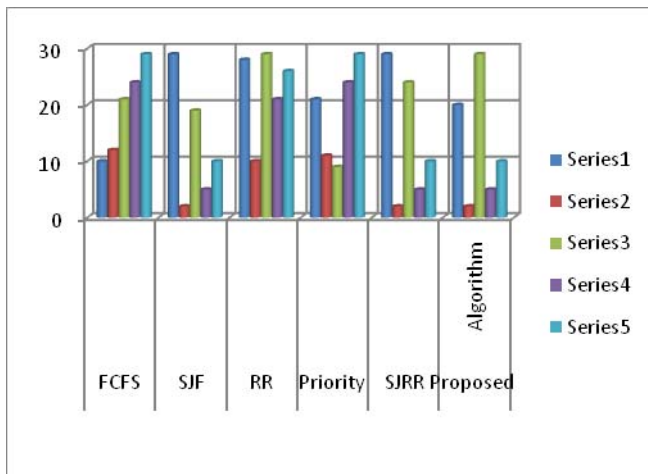


Figure 9: Comparison of Turnaround Time of Proposed Algorithm with Turnaround Time of Existing Algorithms

From the Gantt charts of proposed algorithm and existing algorithms (Figure 2 to 7), it is noticed that waiting time, turnaround time and response time of the proposed algorithms are smaller than existing algorithms. The above two graphs in Figure 8 and 9 also represents that proposed scheduling algorithm is optimum as compared to other existing scheduling algorithms. Maximum CPU utilization and throughput can also be obtained from proposed scheduling algorithm.

From the comparison of the obtained results, it is observed that proposed algorithm has successfully beaten the existing CPU scheduling algorithms. It provides good performance in the scheduling policy. As SJF is optimal scheduling algorithm but for large processes, it gives increased waiting time and sometimes long processes will never execute and remain starved. This problem can be overcome by the proposed algorithm. In future, the working of proposed algorithm will be tested on any open source operating system.

REFERENCES

- [1] Abraham Silberschatz, Peter Baer Galvin, Greg Gagne, "Operating System Concepts", Sixth Edition.
- [2] Andrew S. Tanenbaum, Albert S. Woodhull, "Operating Systems Design and Implementation", Second Edition
- [3] Mohammed A. F. Husainy, "Best-Job-First CPU Scheduling Algorithm", Information Technology Journal 6(2): 288-293, 2007, ISSN 1812-5638
- [4] E. O. Oyetunji, A. E. Oluleye, "Performance Assessment of Some CPU Scheduling Algorithms", Research Journal of Information Technology 1(1): 22-26, 2009, ISSN: 2041-3114
- [5] Sindhu M., Rajkamal R., Vigneshwaran P., "An Optimum Multilevel CPU Scheduling Algorithm," ACE, pp.90-94, 2010 IEEE International Conference on Advances in Computer Engineering, 2010
- [6] Milan Milenkovic, "Operating System Concepts and Design", McGRAW-HILL, Computer Science Series, Second Edition.
- [7] Saeeda Bibi, Farooque Azam, Sameera Amjad, Wasi Haider Butt, Hina Gull, Rashid Ahmed, Yasir Chaudhry "An Efficient SJRR CPU Scheduling Algorithm" International Journal of Computer Science and Information Security, Vol. 8, No. 2, 2010
- [8] Maj. Umar Saleem Butt and Dr. Muhammad Younus Javed, "Simulation of CPU Scheduling Algorithms", 0-7803-6355-8/00/\$10.00@2000 IEEE.
- [9] Rami J. Matarneh, "Self Adjustment Time Quantum in Round Robin Algorithm Depending on Burst Time of the Now Running Processes", American Journal of Applied Sciences 6(10): 18311-1837, 2009, ISSN 1546-9239
- [10] Gary Nutt, "Operating Systems, A Modern Perspective", Second Edition
- [11] Andrew S. Tanenbaum, Albert S. Woodhull "A Modern Operating System", Second Edition
- [12] Md. Mamunur Rashid and Md. Nasim Adhtar, "A New Multilevel CPU Scheduling Algorithm", Journals of Applied Sciences 6(9): 2036-2039, 2009.

Enterprise Crypto method for Enhanced Security over semantic web

Talal Talib Jameel

Department of Medical Laboratory Sciences, Al Yarmouk University College
Baghdad, Iraq

Abstract— the importance of the semantic web technology for enterprises activities and other business sectors is addressing new patters which demand a security concern among these sectors. The security standard in the semantic web enterprises is a step towards satisfying this demand. Meanwhile, the existing security techniques used for describing security properties of the semantic web that restricts security policy specification and intersection. Furthermore, it's common for enterprises environments to have loosely-coupled components in the security. RSA used widely to in the enterprises applications to secure long keys and the use of up-to-date implementations, but this algorithm unable to provide a high level of security among the enterprise semantic web. However, different researchers unable to identify whether they can interact in a secure manner based on RSA. Hence, this study aimed to design a new encryption model for securing the enterprise semantic web with taking in account the current RSA technique as a main source of this study.

Keywords: Agent systems, RSA, ECC, recommendation method, XML, RDF, OWL, enterprise application.

I. INTRODUCTION

The threats to security are increasing with the emergence of new technologies such as software agents. There have been many attacks in past where malicious agents entered into agent platforms and destroyed other active agents. Most of researchers refer to the real world scenario where malicious agent destroyed the other agents on the platform [7]. It will be very critical to focus on security when agents will be used for mission critical systems [3]. In that scenario, a security leak could cause a big harm especially among the enterprise applications over semantic web [6]. A software agent knows as an important part of semantic web [11]. The agents help to get and understand information from different semantic constructs, for instance ontologies, Resource Description Framework (RDF) and (XML).

Therefore it is important to secure data and other relevant technologies for safe enterprise semantic web. Multi-agent systems are an environment where different agents collaborate to perform a specific task [5]. The interaction leaves agents in a different enterprise semantic web vulnerable state, where malicious agent can enter to the system. For example, a malicious agent can enter in an agent platform and kill an agent that was used to perform sales. After killing that agent, this malicious agent can process the order and send the payment to wrong party [17].

The rest of this paper is organized as follows. Issues of the study are presented in section 2. Section 3

presents the proposed model. The Expected benefits are presented in section 4. Conclusion also introduced in section 5 followed by the references.

II. ISSUES OF THE STUDY

Often there has been a need to protect information from 'prying eyes'. Moreover, enterprises applications always require a high level of security. There exist several techniques and frameworks for agents' communication, among enterprise semantic web, but none of those provide cross-platform security [1]. For instance, to encrypt data communication between agents. In their technique both source and destination platforms must have a same cryptography algorithm. Most of these approaches negatively affect the performance agent's communication. There are a number of users around the globe using the semantic web applications and a number of agents are created by those users [1]. Therefore, to reduce the bottlenecks, an ad-hoc based authentication is required for agent communication.

A. Enterprise Semantic Applications

The enterprise semantic applications defined as platform-independent for supporting semantic web application which written in different programming languages [8] [11]. The semantic web platform consists of a set of services and protocols that provide the functionality for developing multitiered.

The main enterprise semantic web application features can be addressed into the following:

- Working together with the HTML based application that consists on RDF, OWL, and XML to build the HTML web relation or other formatted data for the client.
- Provide external storage platforms' that are transparent to the author.
- Provide database connectivity, for managing and classifying the data contents.

These technologies are the important constituents of semantic web services. It is therefore very likely that these services will be agent based in the near future. The success of enterprise application will highly rely on the implementation and usage of these web services [16]. Agents can use intelligent collaborations in order to achieve global optimization while adhering to local requirements.

Figure 1 presents the enterprise communication network among its components.

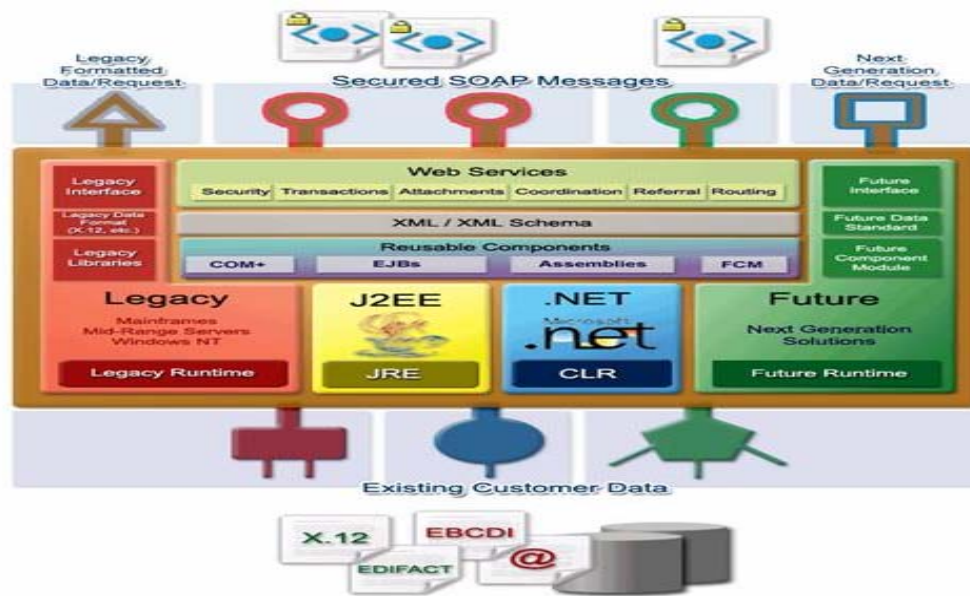


Fig 1. Enterprise communication network

B. Encryption over Semantic Web

Generally, several methods can be used to encrypt data streams, all of which can easily be implemented through software, but not so easily decrypted when either the original or its encrypted data stream are unavailable [13]. (When both source and encrypted data are available, code breaking becomes much simpler, though it is not necessarily easy). The best encryption methods have little effect on system performance, and may contain other benefits (such as data compression) built in.

The current adopting of the new technology have brought a new ideal integration for securing and simplifying the data sharing for all components of enterprise applications [9]. The elements of enterprise application which can be possibly configured within slandered Crypto methods, Table 1 stated the Crypto algorithms comparison:

Table 1. Crypto algorithms comparison [14]

Parameter/algorithm	RSA	ECC	XTR
Key length (bits)	1024	161	Comparable with ECC
Key generation time (processor clocks)	1 261 261 261	40 540 540,5	Less than ECC
Encryption time (processor clocks)	11 261 261,3	3 243 243 243	Comparable with ECC

C. RSA over Semantic Web

Because of the need to ensure that only those eyes intended to view sensitive information can ever see this information, and to ensure that the information arrives unaltered, security systems have often been employed in computer systems for governments, corporations, and

even individuals [18]. Encryption schemes can be broken, but making them as hard as possible to break is the job of a good cipher designer. Figure 2 presents the RSA security process from client to server. As shown, the encrypted client data requested public key from the web decrypts using private key over the internet [15].

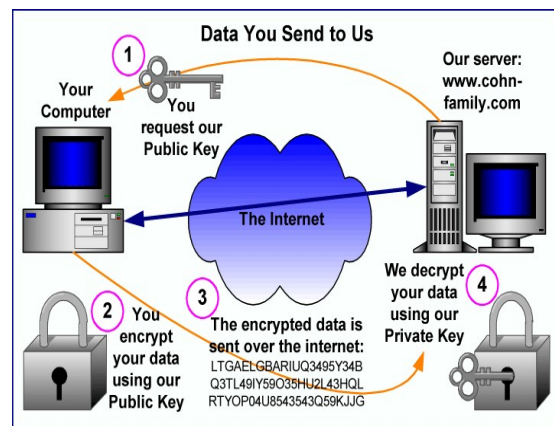


Fig 2. The RSA security over semantic web

This process (encryption) happens when client requests private key from server user name and password. In this way everything client type in and click on can only be decrypted by server through private key.

- 1>> $n = pq$, where p and q are distinct primes.
- 2>> $\phi = (p-1)(q-1)$
- 3>> $e < n$ such that $\gcd(e, \phi) = 1$
- 4>> $d = e^{-1} \mod \phi$.
- 5>> $c = me \mod n, 1 < m < n$.
- 6>> $m = cd \mod n$.

RSA Crypto example

III. THE PROPOSED MODEL

As known, the representing and accessing of the web contents among platforms are determined to be a more recent innovation; most of this representation involves the use of other techniques such as (RDF, XML, and OWL) these technologies works together to link systems together. Enterprise application platform independent facing several security problems in data sharing and accessing which enable web services to work across low level of security. However, the communication process in these platforms (Enterprise application) from the client to the service uses certain technology that helps to translate the client data and assign its security level based XML as the common language. This allows one application to call the services of another application over the network by sending an XML message to it.

Thus, our proposed model will be more efficient in a way that there is no need for agents communication by encrypting the client requests into public store, which reduces the processing and communication time. Also our proposed model will be platform independent because there is no need to maintain standards for cross-platform agents' communication security.

In a pervasive environment, trust can be used for collaboration among devices. Trust can be computed automatically without user interference on the basis of direct and indirect communication [2]. In the direct communication or observation mode the device user's interaction history is considered. For this purpose a trust value is assigned to each identity in the trust database [12]. There exist some formulas such as (Observations and recommendations) that use to calculate the single trust value for the user on the basis of observations and recommendations [2].

This study applies the recommendations technique which aims to specify a degree of trust for each person in the network, for automating trust, which is also called indirect communication [4]. Therefore the observation and recommendation are used together to generate a trust value for a user. Given a user trust value, a trust category is assigned to user with a value of low, medium or high. The trust values should be regularly monitored because when a new recommendation is received new trust value is compared with its old value and trust database is updated by the enterprise application services for single and multi accessing which operate the use access accordingly.

Recommendations are another method of automating trust, which is also called indirect communication [16].

Therefore the observation is used together to generate a trust value for a user. Given a user trust value, a trust category is assigned to user with a value of low, medium or high. The access rights distribution is performed on the basis of the category value. The trust values should be regularly monitored because when a new recommendation is received new trust value is compared with its old value and trust database is updated by update trust category accordingly.

Figure3 and 4 presents the type of trust over enterprise applications which model the logical relationship between the nodes. These nodes will be classified into several groups such as:

- Process Request Group: A request for a service group composed of nodes, node I and node n.
- Register Level Group Provider Group: to provide a service in the network of nodes that comprises the group, as these nodes share certain files, or the provision of certain goods purchases.
- Trust Level Group: trust nodes that comprise the group, node m1, node m2 and node m3.
- Save trust nodes Group: trust network, trust in other nodes on the path formed by the agent.

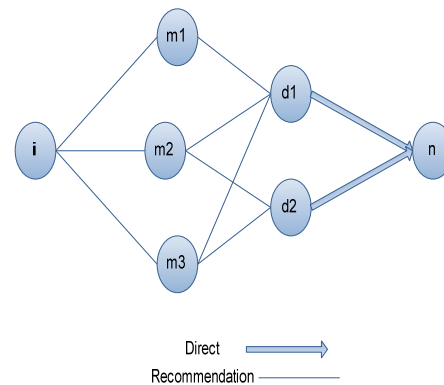


Fig 3. Two type of trust for agent registration level (public store)

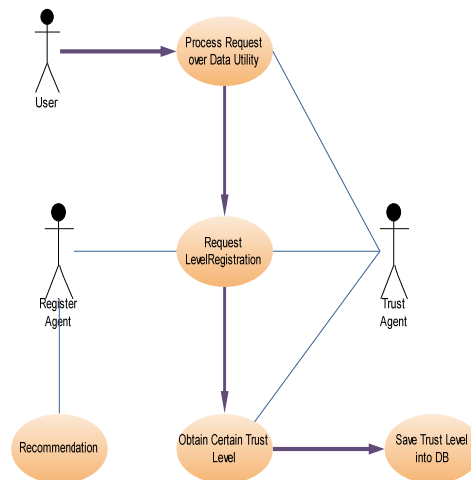


Fig 4. Trust network based recommendation and observation

IV. THE PROPOSED SECURITY MODEL OVER ENTERPRISE SEMANTIC WEB

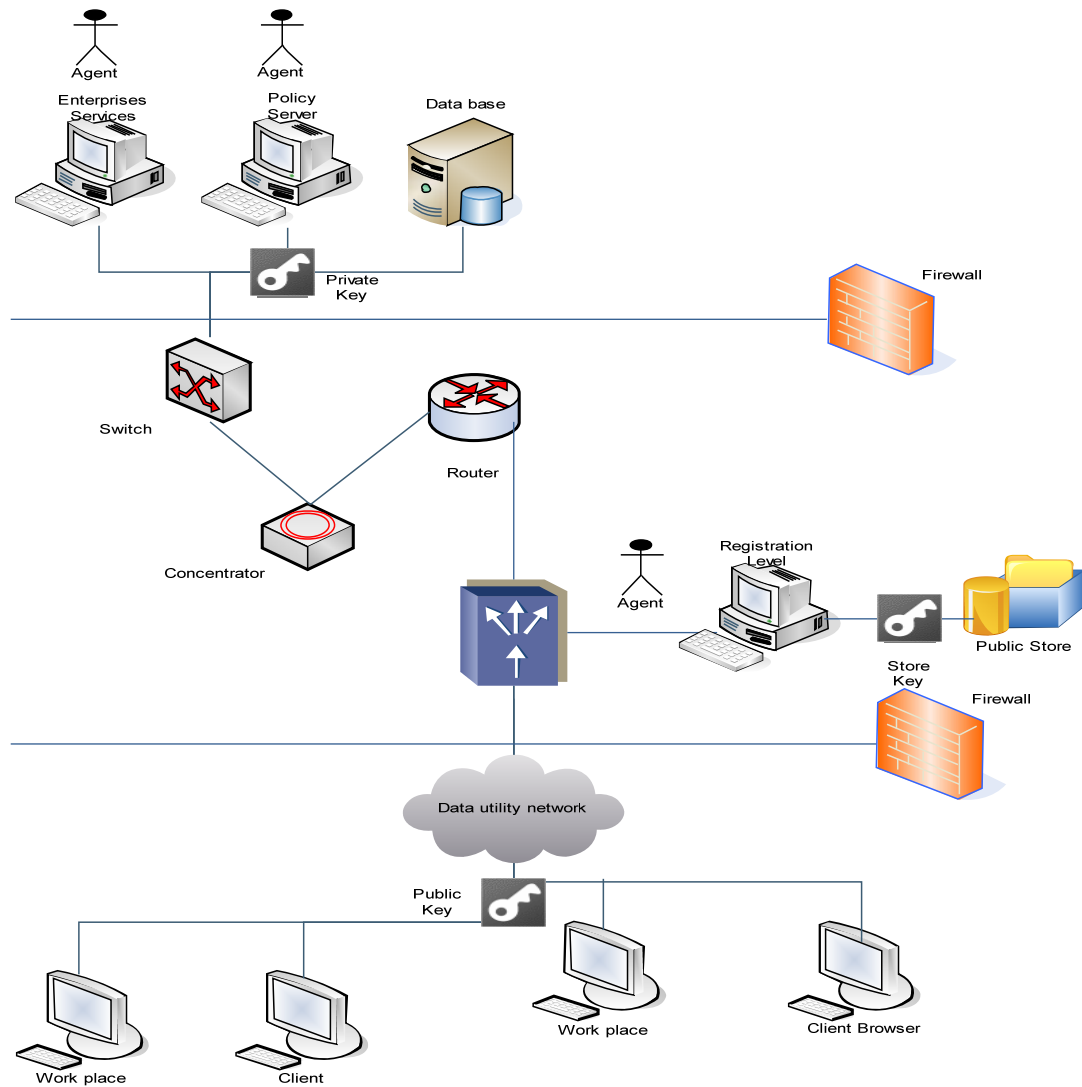


Fig 5. Enterprise Crypto process over semantic web applications

In Figure 5 an agent for registration level outside the environment sends a request to server for registration, server registers it with the lowest security level. With the passage of time the agent becomes more trustworthy based on observations and recommendations.

Delegation is the most important feature in our proposed mechanism through which an agent can delegate set of its rights to another agent for specific period of time. In summary of the whole discussion, we proposed a multi-layered security level mechanism whereby an agent enters in the environment with a low level of security and achieves the higher level of security as it survives in the environment.

V. THE EXPECTED BENEFITS

The expected benefits from the proposed security architecture can be determined the following:

- **Manage user access by level or authority**
this could be done by allowing administrators

or trusted clients to access and share the information across platform based on the retrieved recommendation. Furthermore, this feature will helps to assigns different authorities to different administrators based on specific levels that identified by agent.

- **Determine the Client behavior**
Moreover, the proposed architecture can be capable of customizing the client behaviors based on the security policy contents that over legal clients to use its services and guard against unauthorized use.
- **Provide a High reliability**
Adopting agent systems will helps to simplify the communication performance between client and server.

VI. CONCLUSION

This study aimed to provide a reliable security model for the enterprises semantic web applications based on recommendation method. Meanwhile, the best way for representing and organizing the security for all web resources based platform involves the use of a centralized, identity centric web security system along with a certain language for translating the client request into understandable order based policy enforcement point. Finally, this study was succeeded to determine the working process of the proposed model among web application; also expected benefits were reported in term of Crypto agent technology and recommendation method for assigning the security level for the clients in these applications.

REFERENCES

- [1] V. Bindiganavale, and J. Ouyang, "Role Based Access Control in Enterprise Application Security Administration and User Management, pp.111-117, IEEE, 2006
- [2] M. Youna, and S. Nawaz, "Distributed Trust Based Access Control Architecture to Induce Security in Pervasive Computing" Computer Engineering Department, EME College, NUST Pakistan 2009
- [3] S. Kagal, T. Finin, and Y. Peng, "A Framework for Distributed Trust Management", in Proceedings of IJCAI-01, Workshop on Autonomy, Delegation and Control, Montreal, Canada, 2008
- [4] Foundation for Intelligent Physical Agents. <http://www.fipa.org>
- [5] S. Stojanov, I. Ganchev, I. Popchev, M. O'Droma, and E. Doychev, "An Approach for the Development of Agent- Oriented Distributed eLearning Center," presented at International Conference on Computer Systems and Technologies (CompSysTech), Varna, Bulgaria, 2005
- [6] Y. Li, W. Shen, and H. Ghenniwa, "Agent-Based Web Services Framework and Development Environment," Computational Intelligence, Vol. 20, pp. 678-692, 2004
- [7] J. Hogg, D. Smith, F. Chong, D. Taylor, L. Wall, and P. Slater, "Web service security: Scenarios, patterns, and implementation guidance for Web Services Enhancements (WSE) 3.0," Microsoft Press, March 2006
- [8] M. Schumacher, E. Fernandez-Buglioni, D. Hybertson, F. Buschmann, and P. Sommerlad, "Security patterns: Integrating security and systems engineering," John Wiley and Sons, 2005
- [9] Networked Digital Library of Theses and Dissertations Homepage, <http://www.ndltd.org> (current October 2009)
- [10] Open Archives Initiative Tools, <http://www.openarchives.org/pmh/tools/tools.php> (current October 2009)
- [11] F. Almenarez, A. Marin, C. Campo, and C. Garcia, "TrustAC: Trust-Based Access Control for Pervasive Devices", International conference of Security in pervasive computing, Vol. 3450, pp. 225-238, 2005
- [12] M. Haque, and S. Iqbal, "Security in Pervasive Computing: Current Status and Open Issues", International Journal of Network Security, Vol. 3, No. 3, pp.203-214, 2009
- [13] D. Zuquim, and M. Beatriz, "Web Service Security Management Using Semantic Web Techniques," SAC'08, pp. 2256-2260, Fortaleza, Cear , Brazil, 2008
- [14] W. Abramowicz, A. Ekelhart, S. Fenz, M. Kaczmarek, M. Tjoa, E. Weippl, D. Zyskowski, "Security Aspects in Semantic Web Services Filtering," Proceedings of iiWAS2007, pp. 21-31, Vienna, Austria, 2007
- [15] T. Haytham, M. Koutb, and H. Suoror, "Semantic Web on Scope: A New Architectural Model for the Semantic Web," Journal of Computer Science, Vol. 4 (7): pp.613-624, 2008
- [16] S. Aljawameh, F. Alkhateeb, and E. Maghayreh, "A Semantic Data Validation Service for Web Applications", Journal of Theoretical and Applied Electronic Commerce Research, Vol 5 (1): pp. 39-55, 2010
- [17] D. Sraavan, and M. Upendra, "Privacy for Semantic Web Mining using Advanced DSA Spatial LBS Case Study," (IJCSIS) International Journal on Computer Science and Engineering, Vol. 02, (03): pp. 691-694, 2010
- [18] L. Zheng, and A. Myers, "Securing Nonintrusive Web Encryption through Information Flow," PLAS'08, pp.1-10, Tucson, Arizona, USA. 2008



Ass. L. Mr. Talal Talib Jameel received his Bachelor Degree in statistics from Iraq (1992) and his Master in Information and communication Technology (ICT) from University Utara Malaysia (UUM). Currently, he is working in university of al Yarmouk College as a assistant lecture His research interests in Network Security, Routing Protocols and Electronic learning. He has produced many publications in Journal international rep and also presented papers in International conferences.

On the Performance of Symmetrical and Asymmetrical Encryption for Real-Time Video Conferencing System

Maryam Feily, Salah Noori Saleh, Sureswaran Ramadass
National Advanced IPv6 Centre of Excellence (NAv6)
Universiti Sains Malaysia (USM)
Penang, Malaysia

Abstract— Providing security for video conferencing systems is in fact a challenging issue due to the unique requirements of its real-time multimedia encryption. Modern cryptographic techniques can address the security objectives of multimedia conferencing system. The efficiency of a viable encryption scheme is evaluated using two critical performance metrics: Memory usage, and CPU usage. In this paper, two types of cryptosystems for video conferencing system were tested and evaluated. The first cryptosystem is asymmetric, whereas the second is symmetric. Both cryptosystems were integrated and tested on a commercial based video and multimedia conferencing platform.

Keywords- Encryption; Asymmetric; Symmetric; Security; Efficiency; Video Conferencing.

I. INTRODUCTION

Video and multimedia conferencing systems are currently one of the most popular real-time multimedia applications and have gained acceptance as an Internet based application as well. And since the Internet is involved, security has now become a very important aspect of such systems. To provide a secure video conferencing system, cryptography is used to address data confidentiality and authentication. However, unlike plaintext, encryption of multimedia data, including compressed audio and video, is a challenging process due to the following two constraints. First, the multimedia data encryption and decryption must be done within real-time constraints with minimal delays. Hence, applying heavy encryption algorithms during or after the encoding phase will increase the delay, and are likely to become a performance bottleneck for real-time multimedia applications. The second constraint is that multimedia data is time dependent, and must be well synchronized. Therefore, the needed encryption must be done within the defined time restrictions to keep temporal relations among the video streams intact [1]. There are also other limitations due to the large size of multimedia data [2], [3], but the operation system's network layer can be called upon to handle this. In overall, a viable security mechanism for real-time multimedia transmission must consider both security and efficiency [4].

Since the mid 90's, numerous efforts have been devoted towards the development of real-time multimedia encryption solutions. However, most of the proposed algorithms are characterized by a significant imbalance between security and efficiency. Some of them are efficient enough to meet the requirements of the multimedia encryption, but only provide limited security, whilst others are robust enough to meet the security demands but require complex computations [5].

This paper proposes a viable multimedia encryption that addresses the requirements of video conferencing systems. The efficiency of the proposed encryption scheme is evaluated using two critical performance metrics: Memory usage, and CPU usage. In this paper, the performance of two different types of cryptosystems (symmetric and asymmetric encryption) for encrypting real-time video data are tested and evaluated based on the aforementioned performance metrics. Performance tests of both encryption schemes have been carried out using the *Multimedia Conferencing System (MCS)* [6] that is a commercial video conferencing application.

The first encryption system is an asymmetric cryptosystem based on Elliptic Curve Cryptography (ECC) [7], whereas the second encryption scheme is based on Blowfish [8] which is a symmetric cryptosystem. These schemes have been chosen as the best representative of each symmetric and asymmetric encryption based on their advantages. In fact, ECC is a recent public key cryptosystem which is more efficient and faster than the other asymmetric cryptosystems [9]. On the other hand, Blowfish is known as the fastest symmetric encryption scheme which is compact and suitable for large blocks of data, and therefore suitable for video data encryption [8].

The rest of this paper is organized as follows: Section II provides an overview of cryptographic schemes and compares symmetric and asymmetric cryptography. Section III discusses the asymmetric encryption scheme for real-time video conferencing system, while Section IV discusses the symmetric encryption scheme. Section V provides details on performance tests and a comparison of both cryptosystems. Finally the paper will be concluded in Section VI.

II. OVERVIEW OF CRYPTOGRAPHY

Cryptography is the art and science of hiding secret documents [9]. Security is very important in applications like multimedia conferencing system. To provide a secure multimedia conferencing system, cryptography is used to address data confidentiality, and authentication [10]. Modern cryptographic techniques address the security objectives of multimedia conferencing systems. In general, there are two main categories of cryptography; symmetric and asymmetric key cryptography [9], [11].

A brief overview of each category will be provided in this Section. In addition, symmetric and asymmetric cryptography will be compared briefly to realize the advantages and disadvantages of each one.

A. Symmetric Key Cryptography

Symmetric key cryptography is one of the main categories of cryptography. In symmetric key cryptography, to provide a secure communication a shared secret, called "Secret Key", must be established between sender and recipient. The same key is used for both encryption and decryption. Thus, such a cryptosystem is called "Symmetric" [9]. This type of cryptography can only provide data confidentiality, and cannot address the other objectives of security [9], [11].

Moreover, symmetric key cryptography cannot handle communications in large n -node networks. To provide a confidential communication in a large network of n nodes, each node needs $n-1$ shared secrets. Hence, $n(n-1)$ shared secrets need to be established that is highly impractical and inconvenient for a large value of n [11]. All classical cryptosystems that were developed before 1970s and also most modern cryptosystems are symmetric [11]. DES (Data Encryption Standard) [12], 3DES (Triple Data Encryption Standard) [13], AES (Advanced Encryption Standard) [14], IDEA [15], RC5 [16], Blowfish [8], and SEAL [17] are some of the popular examples of modern symmetric key cryptosystems.

Amongst all symmetric encryption schemes, Blowfish [8] is known as the fastest symmetric encryption scheme which is compact and suitable for large blocks of data, and therefore suitable for video data encryption [8]. Thus, Blowfish is chosen as the best example of symmetric scheme for video encryption in this research.

B. Asymmetric Key Cryptography

Asymmetric or public key cryptography is the other category of cryptography. Despite symmetric key cryptography, public key cryptosystems use a pair of keys instead of a single key for encryption and decryption. One of the keys, called "Public Key", is publicly known and is distributed to all users, whereas the "Private Key" must be kept secret by the owner. Data encrypted with a specific public key, can only be decrypted using the corresponding private key, and vice versa. Since different keys are used for encryption and decryption, the cryptosystem is called "Asymmetric" [9].

Modern public key cryptosystems rely on some computationally intractable problems, and the security of public key cryptosystems depends on the difficulty of the hard problem on which they rely. Hence, public key algorithms operate on sufficiently large numbers to make the cryptanalysis practically infeasible, and thus make the system secure [9], [18]. However, due to smart modern cryptanalysis and modern high speed processing power, the key size of public key cryptosystems grew very large [11]. Using large keys is one of the disadvantages of public key cryptography due to the large memory capacity and large computational power required for key processing.

There are several standard public key algorithms such as RSA [19], El-Gamal [20] and Elliptic Curve Cryptography (ECC) [7]. However, ECC [7] is a recent public key cryptography which is more efficient and faster than the other asymmetric cryptosystems. Unlike previous cryptography solutions, ECC is based on geometric instead of number theory [9]. In fact, the security strength of the ECC relies on the Elliptic Curve Discrete Logarithm Problem (ECDLP) applied to a specific point on an elliptic curve [21], [22]. In ECC, the private key is a random number, whereas the public key is a point on the elliptic curve which is obtained by multiplying the private key with the generator point G on the curve [18]. Hence, computing public key from private key is relatively easy, whereas obtaining private key from public key is computationally infeasible. This is considered as ECDLP that is much more complex than the DLP, and it is believed to be harder than integer factorization problem [18]. Hence, ECC is one of the strongest public key cryptographic systems known today.

In addition, ECC uses smaller keys than the other public key cryptosystems, and requires less computation to provide a high level of security. In other words, efficiency is the most important advantage of the ECC since it offers the highest cryptographic strength per bit [9], [23]. This is a great advantage in many applications, especially in cases that the computational power, bandwidth, storage and efficiency are critical factors [9], [23]. Thus, ECC has been chosen as the best asymmetric encryption in this research.

C. Symmetric Versus Asymmetric Key Cryptography

Despite the Public key cryptography that can only provide data confidentiality, asymmetric key cryptography addresses both data confidentiality and authentication. Public key cryptography solves the problem of confidential communication in large n -node networks, since there is no need to establish a shared secret between communicating parties. Moreover, there are protocols that combine public key cryptography, public key certificates and secure hash functions to enable authentication [11].

However, public key cryptosystems are significantly slower than symmetric cryptosystems. Moreover, public key cryptography is more expensive since it requires large memory capacity and large computational power. For instance, a 128-bit key used with DES provides approximately the same level of security as the 1024-bit key used with RSA [24]. A brief comparison of symmetric and asymmetric key cryptography is summarized in Table I.

TABLE I. SYMMETRIC VERSUS ASYMMETRIC CRYPTOGRAPHY

Cryptosystem	Symmetric	Asymmetric
Confidentiality	Yes	Yes
Data Integrity	No	Yes
Authentication	No	Yes
Number of Keys	1	2
Key Size	Smaller	Larger
Speed	Faster	Slower
Memory Usage	Less	More
Computational Overhead	Less	More
Good for N-node Networks	No	Yes
Some Examples	DES/RC5/Blowfish	RSA/El-Gamal/ECC

III. ASYMMETRIC ENCRYPTION FOR VIDEO CONFERENCING

The asymmetric cryptosystem [25] based on ECC [7] will be reviewed in this Section. In addition, this Section will describe how this encryption scheme was implemented into the MCS video conferencing system.

A. ECC-Based Cryptosystem

The asymmetrical encryption scheme that is tested in this research is a public key cryptosystem based on the Elliptic Curve Digital Signature Algorithm (ECDSA) [25]. It is a robust security platform that employs the most advanced algorithms recognized by the global cryptography community to meet the severe security requirements of certain applications. Furthermore, it is a multilayer cryptosystem which consists of multi layers of public-private key pairs [25]. In its standard mode of encryption, this cryptosystem only uses *256-bit* ECC to encrypt the data. Although this cryptosystem is an ECC public key cryptosystem, it uses other encryption algorithms as well. Mainly, it uses ECDSA for authentication, AES and RSA for key encryption and SHA-2 for hashing.

However, since this cryptosystem is based on ECDSA, the security strength of its encryption scheme mostly relies on the Elliptic Curve Discrete Logarithm Problem (ECDLP) applied to a specific point on an elliptic curve. Hence, breaking this cryptosystem is theoretically equivalent to solving ECDLP, which is computationally impractical for a large key size of *256-bit* [25].

B. Implementation of Asymmetric Scheme

As mentioned earlier, a proper security solution for video conferencing system must address authentication and data confidentiality [9]. However, authentication is well addressed by most video conference systems. Therefore, in order to have a secure video conferencing system, data confidentiality must be provided. Thus, in this research, the aforementioned asymmetric encryption [25] is applied only to the video component of the MCS [6] to protect the video stream. There are two modules in video component responsible for video encryption and decryption that are "Video Capture" and "Video Playback" correspondingly. The architecture of Video Capture and Video Playback are depicted in Fig. 1 and Fig. 2 respectively.

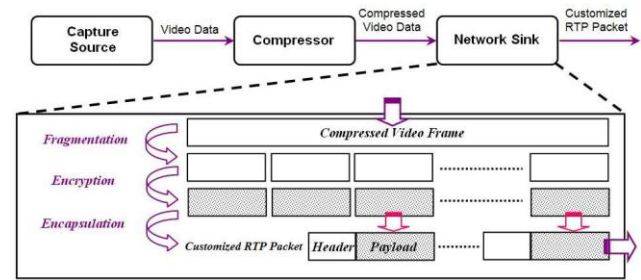


Figure 1. Video Capture Architecture

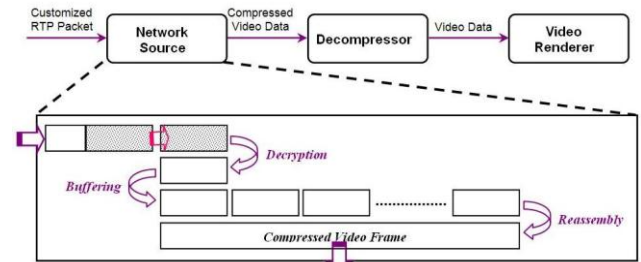


Figure 2. Video Playback Architecture

In addition, it is important to mention that all encryptions and decryptions are performed only at the clients. In this architecture, video encryption and decryption are both performed within the application layer.

After integration of the ECC-based cryptosystem [25] into the video component of the MCS [6], the performance of the system was tested to evaluate the efficiency of asymmetric encryption for real-time video data. The result and analysis of the performance test are presented in Section V.

IV. SYMMETRIC ENCRYPTION FOR VIDEO CONFERENCING

In this Section, an alternative symmetric cryptosystem scheme for video conferencing system is discussed. Amongst all known symmetric encryption such as DES [12], 3DES [13], AES [14], IDEA [15], and RC5 [16], using Blowfish [8] for video data encryption is suggested as it is known to be a fast and compact encryption suitable for large blocks of data [8]. The symmetrical encryption scheme based on Blowfish was implemented by using OpenVPN [26], [27]. In this Section, Blowfish encryption is introduced, and the algorithm is explained briefly. Furthermore, the details of implementing this security scheme into the MCS are explained.

A. Blowfish Encryption

Blowfish is a symmetric block cipher based on the Feistel network. The block size is *64 bits*, whereas the key can be any length up to *448 bits*. Blowfish algorithm consists of two phases: Key Expansion and Data Encryption [8].

In Key Expansion phase a key of at most *448 bits* will be converted into several subkey arrays with maximum of *4168 bytes* which will be used in the Data Encryption phase afterward. During the encryption phase, blocks of *64-bit* input data will be encrypted using a *16-round* Feistel network. Each round of this algorithm consists of permutations and

substitutions. Permutations are key dependant, whereas substitutions depend on both key and data. Decryption is exactly the same as encryption, except that subkeys are used in the reverse order. All operations are XORs and additions on 32-bit words. In addition to these operations, there are four indexed array data lookups per round. Ultimately, the algorithm is cost-effective due to its simple encryption function. Moreover, Blowfish is the fastest block cipher available [8]. Table II shows the speed comparison of block ciphers on a Pentium based computer [8].

TABLE II. SPEED COMPARISON OF BLOCK CIPHERS ON A PENTIUM

Algorithm	Number of Clock Cycles Per Round	Number of Rounds	Number of Clock Cycles per Byte Encrypted
Blowfish	9	16	18
Khufu	5	32	20
RC5	12	16	23
DES	18	16	45
IDEA	50	8	50
Triple DES	18	48	108

B. Implementation of Symmetric Scheme

In order to implement the symmetrical encryption scheme based on Blowfish, OpenVPN software [26] is used as it provides the advantage of choosing from a wide range of cryptographic algorithms according to the level of security required. OpenVPN's cryptography library implements a broad range of standard algorithms to efficiently address both data confidentiality and authentication [26], [27].

For implementation, a VPN server is installed and configured to run in UDP and SSL (Secure Socket Layer) mode as the MCS uses UDP for its video stream, and the SSL Mode is more scalable than the Static Key Mode [27]. Most importantly, Blowfish CBC-mode with 128-bit is selected as the symmetric cipher for data channel encryption to implement the alternative symmetric encryption scheme. In order to provide a multi layer encryption equal to the first scheme, SHA1 with 160-bit message digest is chosen as the hash function algorithm, and 1024-bit RSA as the asymmetric cipher for the control channel to provide authentication. The implemented VPN tunneling and secure data transmission scheme is illustrated in Fig. 3 below.

In this scheme, VPN implements a reliable transport layer on top of UDP using SSL/TLS standard protocol. In other words, a secure layer is established between transport layer and application layer. Hence, it provides a highly secure and reliable connection without the implementation complexities of the network level VPN protocols.

The performance of this scheme is tested on the commercial conferencing system, MCS [6] to realize the efficiency of Blowfish as a symmetric encryption for real-time video data. The results of the performance test and evaluation are presented in Section V.

V. PERFORMANCE TEST AND EVALUATION

In this Section, the performance test and evaluation of both symmetrical and asymmetrical encryption schemes for video conferencing are explained in details, and a comparison of both schemes is provided. In fact, the performance of both encryption schemes is tested to evaluate the efficiency of each scheme and to choose the optimal encryption scheme for real-time video conferencing system.

A. Performance Test

Performance tests of both symmetric and asymmetric encryption schemes have been carried out on the MCS [6] that is a commercial conferencing application. In order to test and evaluate the performance of these cryptosystems, two critical performance parameters namely, the average of CPU usage and the average of Memory usage were measured. These parameters are then compared with a baseline that is the performance of the video conferencing system without any video data encryption/decryption. However, it is important to mention that both encryption schemes have been tested and evaluated only in terms of efficiency, but not security; since

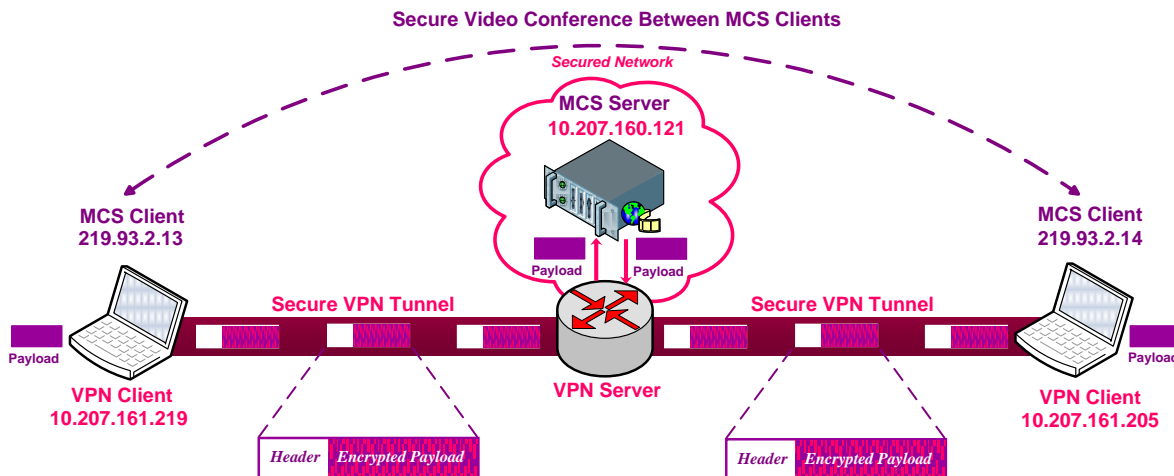


Figure 3. VPN Tunneling and Secure Data Transmission Scheme

the security strength of both encryption schemes are confirmed [8], [25].

All testing have been performed on the same test bed, using identical clients with the following configuration in Table III. This is the recommended system specification for a typical video conference client using the MCS.

First, to provide a baseline for performance evaluation, the performance of the MCS without any video encryption/decryption is tested, and intended parameters are measured. The measurement test bed comprised a video conference between two clients connected to the LAN with a high speed network connection with the speed of *100 Mbps*. At the next stage, the same critical parameters have been measured after applying each encryption schemes. Testing of each case was performed for 80 sessions of video conference between two clients using the MCS, and the average of intended parameters (Memory usage, and CPU usage) was calculated.

B. Evaluation of Performance Result

In this part, the performance results of both symmetric and asymmetric cryptosystems are compared to evaluate the efficiency of each scheme, and to choose the appropriate encryption scheme for real-time video conferencing. The results of CPU usage and Memory usage of both schemes are depicted in Fig. 4 and Fig. 5 respectively.

According to the results, applying asymmetric encryption [25] to the video component increases both CPU usage and Memory usage significantly. The noticeable increase of the CPU usage shown in Fig. 4 is related to the Video Capture module, and shows the heavy processing of the *256-bit* ECC-based encryption. Moreover, as it is illustrated in Fig. 5, the Memory usage is also high and it keeps increasing during the video conference. This is due to the excess Memory usage by the cryptosystem as it creates several memories to encrypt each block of raw data. The dramatic increase of CPU usage and Memory usage are considered as performance bottleneck for the video conferencing system due to the limited processing power and memory capacity.

In contrast, the symmetrical encryption based on Blowfish [8] is more cost-effective in terms of both CPU and Memory usage. Fig. 4 shows that applying symmetric encryption for video conferencing increases the average CPU usage slightly. The 2% increase of the CPU usage is due to the Blowfish encryption and decryption which is obviously far less than the CPU usage of the *256-bit* ECC-based encryption.

It is important to mention that OpenVPN [26] that is used to implement asymmetric encryption uses public key cryptography only for authentication which is mainly done in

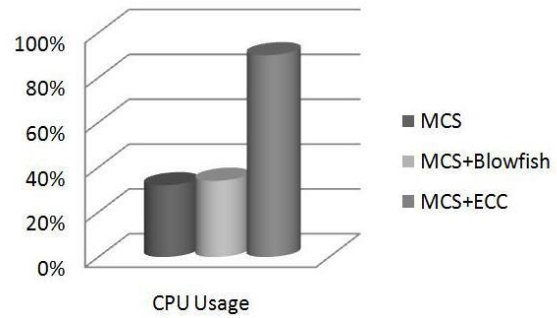


Figure 4. Comparison of CPU Usage

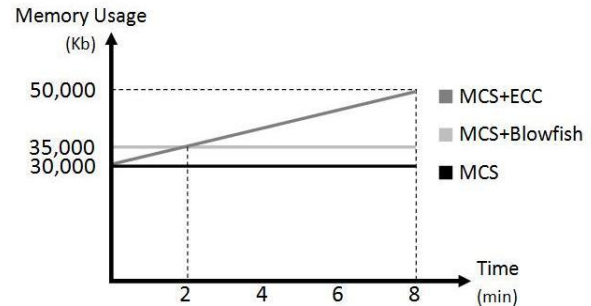


Figure 5. Comparison of Memory Usage

the VPN server, and does not affect the CPU usage of the clients. Moreover, unlike ECC-based encryption, Blowfish cipher does not require a large amount of memory, since it is a compact cipher with a small key size of *128-bit* [8]. In addition, Blowfish encrypts and decrypts the payload of each UDP packet, without creating any memory. Therefore, Memory usage grows by almost a fixed amount of *5000 Kb* as shown in Fig. 5. However, the slight increase in CPU usage and Memory usage is acceptable and does not affect the overall performance of video conferencing system.

VI. CONCLUSION AND FUTURE WORK

In this paper, the performance of two different encryption schemes for real-time video encryption for video conferencing is evaluated in terms of efficiency. The first encryption was an asymmetric cryptosystem based on Elliptic Curve Cryptography (ECC), whereas the second cryptosystem was an alternative symmetric encryption based on Blowfish cipher. These schemes have been chosen as the best representative of each symmetric and asymmetric encryption based on their advantages. Performance tests of both encryption schemes have been carried out on the MCS [6] that is a commercial application. According to the results, the ECC-based cryptosystem [25] caused significant performance bottleneck, and was not effective for real-time video encryption. In contrast, the alternative symmetric encryption based on Blowfish cipher [8] worked well with the MCS [6], and proved to be efficient for encrypting video data in real-time as it is capable to provide an acceptable balance between efficiency and security demands of video and multimedia conferencing systems.

TABLE III. SYSTEM SPECIFICATION OF CLIENTS

Platform	Windows XP Professional (SP2)
Processor	P4 1.80 GHz
RAM	512MB
Hard Disk	40 GB

Performance analysis shows that the inefficiencies of the ECC-based encryption [25] are in fact due to the expensive and heavy computation of the underlying cryptosystem which is a multi layer public key encryption. In fact, ECC public key is suitable to address authentication, and it is not proper for real-time video encryption. However, authentication is usually well addressed by most video conferencing systems, and just a proper encryption for real-time video data is required. Hence, ECC-based encryption is not appropriate for real-time video conferencing as it fails to provide an acceptable balance between efficiency and security demands of video conference. Yet, it is a robust security solution either for non real-time applications or instant messaging where the data is an ordinary text, but not a huge video stream. Unlike ECC-based cryptosystem that sacrifices efficiency for security, the symmetric encryption based on Blowfish meets both security demands and real-time requirements of the video conferencing system with a better performance. It is concluded that the Blowfish which is known as the fastest block cipher is the optimal scheme for real-time video encryption in video conferencing systems.

Nevertheless, there are also few drawbacks of the symmetric encryption scheme implementation using OpenVPN. First, if the VPN server and the video conference server are not located in a secure network, the transmission is not totally secure. Moreover, there will be the problem of single point of failure due to the central VPN server. Hence, the first idea for future work is to implement VPN server directly into the video conference server to eliminate these problems. However, during the time, there will be definitely other new ideas and requirements for the future.

ACKNOWLEDGMENT

The authors graciously acknowledge the support from the *Universiti Sains Malaysia (USM)* through the USM Fellowship awarded to Maryam Feily.

REFERENCES

- [1] Hosseini, H.M.M., Tan, P.M.: Encryption of MPEG Video Streams. In: 2006 IEEE Region 10 Conference (TENCON 2006), pp. 1 - 4. IEEE Press (2006).
- [2] Wu, M.Y., Ma, S., Shu, W.: Scheduled video delivery-a scalable on-demand video delivery scheme. J. IEEE Transactions on Multimedia. 8, 179- -187 (2006).
- [3] Zeng, W., Zhuang, X., Lan, J.: Network friendly media security: rationales, solutions, and open issues. In: IEEE International Conference on Image Processing (2004).
- [4] Choo, E. et.al.: SRMT: A lightweight encryption scheme for secure real-time multimedia transmission. In: IEEE International Conference on Multimedia and Ubiquitous Engineering (MUE'07), pp. 60- -65. IEEE Press (2007).
- [5] Liu, F., Koenig, H.: A novel encryption algorithm for high resolution video. In: ACM International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV'05), pp. 69- -74. ACM New York (2005).
- [6] MLABS.Sdn.Bhd: Multimedia Conferencing System - MCS Ver.6 Technical White paper (2005). Available at <http://www.mlabs.com/paper/MCSv6.pdf>.
- [7] Certicom: SEC 1: Elliptic Curve Cryptography. Vol. 1.5 1.0, 2005. Available at http://www.secg.org/download/aid-385/sec1_final.pdf.

- [8] Schneier, B.: Description of a New Variable-Length Key, 64-Bit Block Cipher (Blowfish). In: Fast Software Encryption, Cambridge Security Workshop (December 1993), pp. 191- -204. Springer-Verlag (1994). Available at <http://www.schneier.com/paper-blowfish-fse.html>.
- [9] Stallings, W.: Cryptography and network security: principles and practice. Prentice Hall (2006).
- [10] Ahmet, M. E.: Protecting Intellectual Property in Digital Multimedia Networks. J. IEEE Computer Society. 36, 39- -45(2003).
- [11] Furht, B., Kirovski, D.: Multimedia Security Handbook. CRC Press LLC (2004).
- [12] National Bureau of Standards.: Data Encryption Standard. National Bureau of Standards, US Department of Commerce- Federal Information Processing, Standards Publication 46(1977).
- [13] Institute, A.N.S.: Triple Data Encryption Algorithm Modes of Operation. American National Standards Institute, ANSI X9.52-1998 (1998).
- [14] Daemen, J., Rijmen, V.: AES proposal: Rijndael (1999). Available at <http://www.nist.gov/CryptoToolkit>.
- [15] Lai, X., Massey, J.L.: A proposal for a new block encryption standard. J. Springer. 90, 389- -404 (1990).
- [16] Rivest, R.: The RC5 encryption algorithm. J. Springer. pp. 86- -96 (1994).
- [17] Rogaway, P., Coppersmith, D.: A software-optimized encryption algorithm. J. Cryptology. 11, 273- -287 (1998).
- [18] Anoop, M.S.: Public key Cryptography: Applications Algorithms and Mathematical Explanations. Tata Elxsi Ltd, India (2007).
- [19] Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. J. Communications of the ACM (1978).
- [20] ElGamal, T.: A Public Key Cryptosystem and A signature Scheme Based on Discrete Logarithm Problem. J. IEEE Transaction on Information Theory. 31, 469- -472 (1985).
- [21] Koblitz, N.: Introduction to Elliptic Curves and Modular Form. Springer-Verlag (1993).
- [22] Miller, V.: Uses of Elliptic Curves in Cryptography. Advances in Cryptology (CRYPTO '85). LNCS, vol. 218, pp. 417- -426. Springer-Verlag (1986).
- [23] Johnson, D. B.: ECC: Future Resiliency, and High Security Systems. In: Certicom PKS '99 (1999).
- [24] Menezes, A. J., Van Oorschot, P. C., Vanstone, S. A.: Handbook of Applied Cryptography. CRC Press Inc. (1997).
- [25] Zeetoo(M) Sdn. Bhd: Zeetoo Encryptor ECDSA (2006). Available at <http://mymall.netbuilder.com.my/?domain=zeetoo&doit=showclass&cid=6>.
- [26] OpenVPN Technologies Inc.: OpenVPN (2007). Available at <http://www.openvpn.net>.
- [27] Feilner, M.: OpenVPN: Building and Integrating Virtual Private Networks - PACKT Publishing (2006).

AUTHORS PROFILE



Maryam Feily is a Ph.D. Student and a Research Fellow at the *Universiti Sains Malaysia (USM)*. She received the B.Eng. degree in Software Engineering from the Azad University (Iran) in 2002, and the M.Sc. degree in Computer Science from USM (Malaysia) in 2008. She has been awarded with the USM Fellowship in 2009. Furthermore, she is proud of being one of the successful graduates of Iran's National Organization for Development of Exceptional Talents (NODET). Her research interests include Network Management, Network Security, Cyber Security, and Overlay Networks.



Salah Noori Saleh is a Senior Developer and Researcher in the Universiti Sains Malaysia (USM). He has received the Ph.D. degree from USM in 2010. He received the B.Sc. degree in Computer Engineering from the University of Baghdad (Iraq) and the M.Sc. degree in Computer Science from USM (Malaysia). His research interests include Network Architectures and Protocols, Multimedia and Peer-to-Peer Communications, Overlay Networks, and Network Security.



Sureswaran Ramadass is a Professor with the Universiti Sains Malaysia (USM). He is also the Director of the National Advanced IPv6 Centre of Excellence (NAV6) at USM. He received the B.Sc. degree and the M.Sc. degree in Electrical and Computer Engineering from the University of Miami in 1987 and 1990 respectively. He received the Ph.D. degree from the Universiti Sains Malaysia (USM) in 2000 while serving as a full time faculty in the School of Computer Sciences. He is a Primary Member of APAN as well as the Head of APAN Malaysia (Asia Pacific Advanced Networks). He is currently the IPv6 Domain Head for MYREN (Malaysian Research and Education Network) and the Chairman of the Asia Pacific IPv6 Task Force (APV6TF).

RACHSU Algorithm based Handwritten Tamil Script Recognition

C.Sureshkumar

Department of Information Technology,
J.K.K.Nataraja College of Engineering,
Namakkal, Tamilnadu, India.

Dr.T.Ravichandran

Department of Computer Science & Engineering,
Hindustan Institute of Technology,
Coimbatore, Tamilnadu, India

Abstract- Handwritten character recognition is a difficult problem due to the great variations of writing styles, different size and orientation angle of the characters. The scanned image is segmented into paragraphs using spatial space detection technique, paragraphs into lines using vertical histogram, lines into words using horizontal histogram, and words into character image glyphs using horizontal histogram. The extracted features considered for recognition are given to Support Vector Machine, Self Organizing Map, RCS, Fuzzy Neural Network and Radial Basis Network. Where the characters are classified using supervised learning algorithm. These classes are mapped onto Unicode for recognition. Then the text is reconstructed using Unicode fonts. This character recognition finds applications in document analysis where the handwritten document can be converted to editable printed document. Structure analysis suggested that the proposed system of RCS with back propagation network is given higher recognition rate.

Keywords - Support Vector, Fuzzy, RCS, Self organizing map, Radial basis function, BPN

I. INTRODUCTION

Hand written Tamil Character recognition refers to the process of conversion of handwritten Tamil character into Unicode Tamil character. Among different branches of handwritten character recognition it is easier to recognize English alphabets and numerals than Tamil characters. Many researchers have also applied the excellent generalization capabilities offered by ANNs to the recognition of characters. Many studies have used fourier descriptors and Back Propagation Networks for classification tasks. Fourier descriptors were used in to recognize handwritten numerals. Neural Network approaches were used to classify tools. There have been only a few attempts in the past to address the recognition of printed or handwritten Tamil Characters. However, less attention had been given to Indian language recognition. Some efforts have been reported in the literature for Tamil scripts. In this work, we propose a recognitionsystem for handwritten Tamil characters. Tamil is a South Indian language spoken widely in TamilNadu in India. Tamil has the longest unbroken literary tradition amongst the Dravidian languages. Tamil is inherited from Brahmi script. The earliest available text is the Tolkaappiyam, a work

describing the language of the classical period. There are several other famous works in Tamil like Kambar Ramayana and Silapathigaram but few supports in Tamil which speaks about the greatness of the language. For example, Thirukural is translated into other languages due to its richness in content. It is a collection of two sentence poems efficiently conveying things in a hidden language called Slaydai in Tamil. Tamil has 12 vowels and 18 consonants. These are combined with each other to yield 216 composite characters and 1 special character (aayutha ezhuthu) counting to a total of $(12+18+216+1)$ 247 characters. Tamil vowels are called uyireluttu (uyir – life, eluttu – letter). The vowels are classified into short (kuril) and long (five of each type) and two diphthongs, /ai/ and /auk/, and three "shortened" (kuril) vowels. The long (nedil) vowels are about twice as long as the short vowels. Tamil consonants are known as meyyeluttu (mey - body, eluttu - letters). The consonants are classified into three categories with six in each category: vallinam - hard, mellinam - soft or Nasal, and itayinam - medium. Unlike most Indian languages, Tamil does not distinguish aspirated and unaspirated consonants. In addition, the voicing of plosives is governed by strict rules in centamil. As commonplace in languages of India, Tamil is characterised by its use of more than one type of coronal consonants. The Unicode Standard is the Universal Character encoding scheme for written characters and text. The Tamil Unicode range is U+0B80 to U+0BFF. The Unicode characters are comprised of 2 bytes in nature.

II. TAMIL CHARACTER RECOGNITION

The schematic block diagram of handwritten Tamil Character Recognition system consists of various stages as shown in figure 1. They are Scanning phase, Preprocessing, Segmentation, Feature Extraction, Classification, Unicode mapping and recognition and output verification.

A. Scanning

A properly printed document is chosen for scanning. It is placed over the scanner. A scanner software is invoked which scans the document. The document is sent to a program that saves it in preferably TIF, JPG or GIF format, so that the image of the document can be obtained when needed.

B. Preprocessing

This is the first step in the processing of scanned image. The scanned image is preprocessed for noise removal. The resultant image is checked for skewing. There are possibilities of image getting skewed with either left or right orientation. Here the image is first brightened and binarized. The function for skew detection checks for an angle of orientation between ± 15 degrees and if detected then a simple image rotation is carried out till the lines match with the true horizontal axis, which produces a skew corrected image.

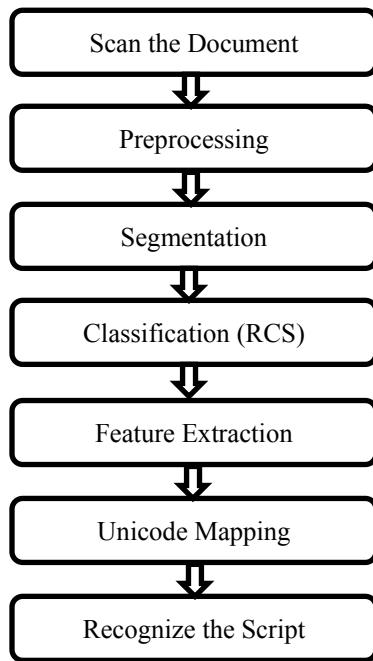


Figure 1. Schematic block diagram of handwritten Tamil Character Recognition system

Knowing the skew of a document is necessary for many document analysis tasks. Calculating projection profiles, for example, requires knowledge of the skew angle of the image to a high precision in order to obtain an accurate result. In practical situations, the exact skew angle of a document is rarely known, as scanning errors, different page layouts, or even deliberate skewing of text can result in misalignment. In order to correct this, it is necessary to accurately determine the skew angle of a document image or of a specific region of the image, and, for this purpose, a number of techniques have been presented in the literature. Figure 1 shows the histograms for skewed and skew corrected images and original character. Postal found that the maximum valued position in the Fourier spectrum of a document image corresponds to the angle of skew. However, this finding was limited to those documents that contained only a single line spacing, thus the peak was strongly localized around a single point. When variant line spacing's are introduced, a series of Fourier spectrum maxima are created in a line that extends from the origin. Also evident is a subdominant line that lies at 90 degrees to the dominant line. This is due to character and word spacing's and the

strength of such a line varies with changes in language and script type. Scholkopf, Simard expand on this method, breaking the document image into a number of small blocks, and calculating the dominant direction of each such block by finding the Fourier spectrum maxima. These maximum values are then combined over all such blocks and a histogram formed. After smoothing, the maximum value of this histogram is chosen as the approximate skew angle. The exact skew angle is then calculated by taking the average of all values within a specified range of this approximate. There is some evidence that this technique is invariant to document layout and will still function even in the presence of images and other noise. The task of smoothing is to remove unnecessary noise present in the image. Spatial filters could be used. To reduce the effect of noise, the image is smoothed using a Gaussian filter. A Gaussian is an ideal filter in the sense that it reduces the magnitude of high spatial frequencies in an image proportional to their frequencies. That is, it reduces magnitude of higher frequencies more. Thresholding is a nonlinear operation that converts a gray scale image into a binary image where the two levels are assigned to pixels that are below or above the specified threshold value. The task of thresholding is to extract the foreground from the background. Global methods apply one threshold to the entire image while local thresholding methods apply different threshold values to different regions of the image. Skeletonization is the process of peeling off a pattern as any pixels as possible without affecting the general shape of the pattern. In other words, after pixels have been peeled off, the pattern should still be recognized. The skeleton hence obtained must be as thin as possible, connected and centered. When these are satisfied the algorithm must stop. A number of thinning algorithms have been proposed and are being used. Here Hilditch's algorithm is used for skeletonization.

C. Segmentation

After preprocessing, the noise free image is passed to the segmentation phase, where the image is decomposed [2] into individual characters. Figure 2 shows the image and various steps in segmentation.

D. Feature extraction

The next phase to segmentation is feature extraction where individual image glyph is considered and extracted for features. Each character glyph is defined by the following attributes: (1) Height of the character. (2) Width of the character. (3) Numbers of horizontal lines present short and long. (4) Numbers of vertical lines present short and long. (5) Numbers of circles present. (6) Numbers of horizontally oriented arcs. (7) Numbers of vertically oriented arcs. (8) Centroid of the image. (9) Position of the various features. (10) Pixels in the various regions.

II. NEURAL NETWORK APPROACHES

The architecture chosen for classification is Support Vector machines, which in turn involves training and testing the use

of Support Vector Machine (SVM) classifiers [1]. SVMs have achieved excellent recognition results in various pattern recognition applications. Also in handwritten character recognition they have been shown to be comparable or even superior to the standard techniques like Bayesian classifiers or multilayer perceptrons. SVMs are discriminative classifiers based on vapnik's structural risk minimization principle. Support Vector Machine (SVM) is a classifier which performs classification tasks by constructing hyper planes in a multidimensional space.

A. Classification SVM Type-1

For this type of SVM, training involves the minimization of the error function:

$$\frac{1}{2} w^T w + c \sum_{i=1}^N \xi_i \quad (1)$$

subject to the constraints:

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0, i = 1, \dots, N \quad (2)$$

Where C is the capacity constant, w is the vector of Coefficients, b a constant and ξ_i are parameters for handling no separable data (inputs). The index i label the N training cases [6, 9]. Note that $y \pm 1$ represents the class labels and x_i is the independent variables. The kernel ϕ is used to transform data from the input (independent) to the feature space. It should be noted that the larger the C, the more the error is penalized.

B. Classification SVM Type-2

In contrast to Classification SVM Type 1, the Classification SVM Type 2 model minimizes the error function:

$$\frac{1}{2} w^T w - \nu \rho + \frac{1}{N} \sum_{i=1}^N \xi_i \quad (3)$$

subject to the constraints:

$$y_i(w^T \phi(x_i) + b) \geq \rho - \xi_i \text{ and } \xi_i \geq 0, i = 1, \dots, N; \rho \geq 0 \quad (4)$$

A self organizing map (SOM) is a type of artificial neural network that is trained using unsupervised learning to produce a low dimensional (typically two dimensional), discredited representation of the input space of the training samples, called a map. Self organizing maps are different than other artificial neural networks in the sense that they use a neighborhood function to preserve the topological properties of the input space.

This makes SOM useful for visualizing low dimensional views of high dimensional data, akin to multidimensional scaling. SOMs operate in two modes: training and mapping. Training builds the map using input examples. It is a competitive process, also called vector quantization [7]. Mapping automatically classifies a new input vector.

The self organizing map consists of components called nodes or neurons. Associated with each node is a

weight vector of the same dimension as the input data vectors and a position in the map space. The usual arrangement of nodes is a regular spacing in a hexagonal or rectangular grid. The self organizing map describes a mapping from a higher dimensional input space to a lower dimensional map space.

C. Algorithm for Kohonon's SOM

(1) Assume output nodes are connected in an array, (2) Assume that the network is fully connected all nodes in input layer are connected to all nodes in output layer. (3) Use the competitive learning algorithm.

$$|\omega_i - x| \leq |\omega_k - x| \quad \forall k \quad (5)$$

$$w_k(\text{new}) = w_k(\text{old}) + \mu \chi(i, k)(x - w_k) \quad (6)$$

Randomly choose an input vector x, Determine the "winning" output node i, where w_i is the weight vector connecting the inputs to output node.

A new neural classification algorithm and Radial- Basis-Function Networks are known to be capable of universal approximation and the output of a RBF network can be related to Bayesian properties. One of the most interesting properties of RBF networks is that they provide intrinsically a very reliable rejection of "completely unknown" patterns at variance from MLP. Furthermore, as the synaptic vectors of the input layer store locations in the problem space, it is possible to provide incremental training by creating a new hidden unit whose input synaptic weight vector will store the new training pattern. The specifics of RBF are firstly that a search tree is associated to a hierarchy of hidden units in order to increase the evaluation speed and secondly we developed several constructive algorithms for building the network and tree.

D. RBF Character Recognition

In our handwritten recognition system the input signal is the pen tip position and 1-bit quantized pressure on the writing surface. Segmentation is performed by building a string of "candidate characters" from the acquired string of strokes [16]. For each stroke of the original data we determine if this stroke does belong to an existing candidate character regarding several criteria such as: overlap, distance and diacriticity. Finally the regularity of the character spacing can also be used in a second pass. In case of text recognition, we found that punctuation needs a dedicated processing due to the fact that the shape of a punctuation mark is usually much less important than its position. it may be decided that the segmentation was wrong and that back tracking on the segmentation with changed decision thresholds is needed. Here, tested two encoding and two classification methods. As the aim of the writer is the written shape and not the writing gesture it is very natural to build an image of what was written and use this image as the input of a classifier.

Both the neural networks and fuzzy systems have some things in common. They can be used for solving a problem (e.g. pattern recognition, regression or density estimation) if there does not exist any mathematical model of

the given problem. They solely do have certain disadvantages and advantages which almost completely disappear by combining both concepts. Neural networks can only come into play if the problem is expressed by a sufficient amount of observed examples [12]. These observations are used to train the black box. On the one hand no prior knowledge about the problem needs to be given. However, it is not straightforward to extract comprehensible rules from the neural network's structure. On the contrary, a fuzzy system demands linguistic rules instead of learning examples as prior knowledge. Furthermore the input and output variables have to be described linguistically. If the knowledge is incomplete, wrong or contradictory, then the fuzzy system must be tuned. Since there is not any formal approach for it, the tuning is performed in a heuristic way. This is usually very time consuming and error prone.

E. Hybrid Fuzzy Neural Network

Hybrid Neuro fuzzy systems are homogeneous and usually resemble neural networks. Here, the fuzzy system is interpreted as special kind of neural network. The advantage of such hybrid NFS is its architecture since both fuzzy system and neural network do not have to communicate any more with each other. They are one fully fused entity [14]. These systems can learn online and offline. The rule base of a fuzzy system is interpreted as a neural network. Thus the optimization of these functions in terms of generalizing the data is very important for fuzzy systems. Neural networks can be used to solve this problem.

F. RACHSU Script Recognition

Once a boundary image is obtained then Fourier descriptors are found. This involves finding the discrete Fourier coefficients $a[k]$ and $b[k]$ for $0 \leq k \leq L-1$, where L is the total number of boundary points found, by applying equations (7) and (8)

$$a[k] = 1/L \sum_{m=1}^L x[m] e^{-jk(2\pi/L)m} \quad (7)$$

$$b[k] = 1/L \sum_{m=1}^L y[m] e^{jk(2\pi/L)m} \quad (8)$$

Where $x[m]$ and $y[m]$ are the x and y co-ordinates respectively of the m^{th} boundary point. In order to derive a set of Fourier descriptors that have the invariant property with respect to rotation and shift, the following operations are defined [3,4]. For each n compute a set of invariant descriptors $r(n)$.

$$r(n) = \left[|a(n)|^2 + |b(n)|^2 \right]^{1/2} \quad (9)$$

It is easy to show that $r(n)$ is invariant to rotation or shift. A further refinement in the derivation of the descriptors is realized if dependence of $r(n)$ on the size of the character is eliminated by computing a new set of descriptors $s(n)$ as

$$s(n) = r(n)/r(1) \quad (10)$$

The Fourier coefficients $a(n)$, $b(n)$ and the invariant descriptors $s(n)$, $n = 1, 2, \dots, (L-1)$ were derived for all of the character specimens [5].

G. RACHSU Algorithm

The major steps of the algorithm are as follows:

1. Initialize all W_{ij} s to small random values with W_{ij} being the value of the connection weight between unit j and unit i in the layer below.
2. Present the 16-dimensional input vector y_0 , input vector consists of eight fourier descriptors and eight border transition values. Specify the desired outputs. If the net is used as a classifier then all desired outputs are typically set to zero except for that corresponding to the class the input is from.
3. Calculate the outputs y_j of all the nodes using the present value of W , where W_{ij} is the value of connection weight between unit j and the unit4 in the layer below:

$$y_i = \frac{1}{1 + \exp(-\sum_j y_j w_{ij})} \quad (11)$$

This particular nonlinear function is called a function sigmoid

4. Adjust weights by :

$$W_{ij}(n+1) = W_{ij}(n) + \alpha \delta_j y_i + \xi (W_{ij}(n) - W_{ij}(n-1))$$

where $0 < \xi < 1$

(12)

where $(n+1)$, (n) and $(n-1)$ index next, present and previous, respectively. The parameter α is a learning rate similar to step size in gradient search algorithms, between 0 and 1 which determines the effect of past weight changes on the current direction of movement in weight space. δ_j is an error term for node j . If node j is an output node, d_j and y_i stand for, respectively, the desired and actual value of a node, then

$$\delta_i = (d_j - y_i) y_i (1 - y_i) \quad (13)$$

If node j is an internal hidden node, then :

$$\delta_j = y_j (1 - y_j) \sum_k \delta_k w_k \quad (14)$$

Where k is over all nodes in the layer above node j .

5. Present another input and go back to step (2). All the training inputs are presented cyclically until weights stabilize (converge).

H. Structure Analysis of RCS

The recognition performance of the RCS will highly depend on the structure of the network and training algorithm. In the proposed system, RCS has been selected to train the network [8]. It has been shown that the algorithm has much better learning rate. Table 1 shows the comparison of various approach classification. The number of nodes in input, hidden and output layers will determine the network structure.

TABLE 1 COMPARISON OF CLASSIFIERS

Type of classifier	Error	Efficiency
--------------------	-------	------------

SVM	0.001	91%
SOM	0.02	88%
FNN	0.06	90%
RBN	0.04	88%
RCS	0	97%

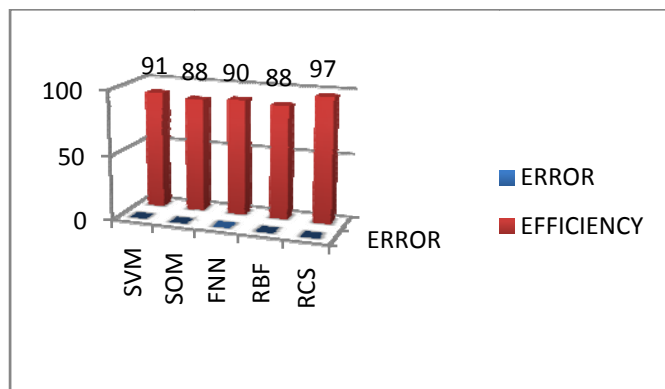


Figure 2 Character Recognition Efficiency and Error report

I.Number of Hidden Layer Nodes

The number of hidden nodes will heavily influence the network performance. Insufficient hidden nodes will cause under fitting where the network cannot recognize the numeral because there are not enough adjustable parameter to model or to map the input output relationship. Figure 2 shows the character recognition efficiency and error report. The minimum number of epochs taken to recognize a character and recognition efficiency of training as well as test character set as the number of hidden nodes is varied. In the proposed system the training set recognition rate is achieved and in the test set the recognized speed for each character is 0.1sec and accuracy is 97%. The training set produced much higher recognition rate than the test set. Structure analysis suggested that RCS is given higher recognition rate.Hence Unicode is chosen as the encoding scheme for the current work. The scanned image is passed through various blocks of functions and finally compared with the recognition details from the mapping table from which corresponding unicodes are accessed and printed using standard Unicode fonts so that the Character Recognition is achieved.

III. EXPERIMENTAL RESULTS

The invariant Fourier descriptors feature is independent of position, size, and orientation. With the combination of RCS and back propagation network, a high accuracy recognition system is realized. The training set consists of the writing samples of 25 users selected at random from the 40, and the test set, of the remaining 15 users. A portion of the training data was also used to test the system. In the training set, a recognition rate of 100% was achieved and in the test set the recognized speed for each character is 0.1sec and accuracy is

97%. Understandably, the training set produced much higher recognition rate than the test set. Structure analysis suggested that RCS with 5 hidden nodes has lower number of epochs as well as higher recognition rate.

IV. CONCLUSION

Character Recognition is aimed at recognizing handwritten Tamil document. The input document is read preprocessed, feature extracted and recognized and the recognized text is displayed in a picture box. The TamilCharacter Recognition is implemented using a Java Neural Network. A complete tool bar is also provided for training, recognizing and editing options. Tamil is an ancient language. Maintaining and getting the contents from and to the books is very difficult. In a way Character Recognition provides a paperless environment. Character Recognition provides knowledge exchange by easier means. If a knowledge base of rich Tamil contents is created, it can be accessed by people of varying categories with ease and comfort.

ACKNOWLEDGEMENT

The researchers would like to thank S. Yasodha and Avantika for his assistance in the data collection and manuscript preparation of this article.

REFERENCES

- [1] B. Heisele, P. Ho, and T. Poggio, "Character Recognition with Support Vector Machines: Global Versus Component Based Approach," in ICCV, 2006, vol. 02, no. 1, pp. 688–694.
- [2] Julie Delon, Agnès Desolneux, "A Nonparametric Approach for Histogram Segmentation," IEEE Trans. On image processing., vol. 16, no. 1, pp. 235–241. 2007
- [3] B. Sachine, P. Manoj, M.Ramya, "Character Segmentations," in Advances in Neural Inf. Proc. Systems, vol. 10. MIT Press, 2005, vol.01, no 02 pp. 610–616.
- [4] O. Chapelle, P. Haffner, and V. Vapnik, "SVMs for Histogram-based Image Classification," IEEE Transactions on Neural Networks, special issue on Support Vectors, vol 05, no 01, pp. 245–252, 2007.
- [5] Simone Marinai, Marco Gori, "Artificial Neural Networks for Document Analysis and Recognition "IEEE Transactions on pattern analysis and machine intelligence, vol.27, no.1, Jan 2005, pp. 652–659.
- [6] M. Anu, N. Viji, and M. Suresh, "Segmentation Using Neural Network," IEEE Trans. Patt. Anal. Mach. Intell., vol. 23, pp. 349–361, 2006.
- [7] B. Scholkopf, P. Simard, A. Smola, and V. Vapnik, "Prior Knowledge in Support Vector Kernels," in Advances in Neural Inf. Proc. Systems, vol. 10. MIT Press, 2007, pp. 640–646.
- [8] Olivier Chapelle, Patrick Haffner, "SOM for Histogram-based Image Classification," IEEE Transactions on Neural Networks, 2005. Vol 14 no 02, pp. 214–230.
- [9] S. Belongie, C. Fowlkes, F. Chung, and J. Malik, "Spectral Partitioning with Indefinite Kernels Using the Nystrom Extention," in ECCV, part III, Copenhagen, Denmark, 2006, vol 12 no 03, pp. 123–132
- [10] T. Evgeniou, M. Pontil, and T. Poggio, "Regularization Networks and Support Vector Machines," Advances in Computational Mathematics, vol. 13, pp. 1–11, 2005.
- [11] P. Bartlett and J. Shawe Taylor, "Generalization performance of support vector machines and other pattern classifiers," in Advances in

- Kernel Methods Support Vector Learning. 2008, MIT Press Cambridge, USA, 2002, vol 11 no 02, pp. 245-252.
- [12] E.Osuna, R.Freund, and F.Girosi, "Training Support Vector machines: an application to face detection," in IEEE CVPR'07, Puerto Rico, vol 05 no 01, pp. 354-360, 2007.
- [13] V. Johari and M. Razavi, "Fuzzy Recognition of Persian Handwritten Digits," in Proc. 1st Iranian Conf. on Machine Vision and Image Processing, Birjand, vol 05 no 03, 2006, pp. 144-151.
- [14] P. K. Simpson, "Fuzzy Min-Max Neural Networks- Part1 Classification," IEEE Trans. Neural Network., vol. 3, no. 5, pp. 776-786, 2002.
- [15] H. R. Boveiri, "Scanned Persian Printed Text Characters Recognition Using Fuzzy-Neural Networks," IEEE Transaction on Image Processing, vol 14, no 06, pp. 541-552, 2009.
- [16] D. Deng, K. P. Chan, and Y. Yu, "Handwritten Chinese character recognition using spatial Gabor filters and self-organizing feature maps", Proc. IEEE Inter. Confer. On Image Processing, vol. 3, pp. 940-944, 2004.

AUTHORS PROFILE

C.Sureshkumar received the M.E. degree in Computer Science and Engineering from K.S.R College of Technology, Thiruchengode, Tamilnadu, India in 2006. He is pursuing the Ph.D degree in Anna University Coimbatore, and going to submit his thesis in Handwritten Tamil Character recognition using Neural Network. Currently working as HOD and Professor in the Department of Information Technology, in JKKN College of Engineering and Technology, Tamil Nadu, India. His current research interest includes document analysis, optical character recognition, pattern recognition and network security. He is a life member of ISTE.

Dr. T. Ravichandran received a Ph.D in Computer Science and Engineering in 2007, from the University of Periyar, Tamilnadu, India. He is working as a Principal at Hindustan Institute of Technology, Coimbatore, Tamilnadu, India, specialised in the field of Computer Science. He published many papers on computer vision applied to automation, motion analysis, image matching, image classification and view-based object recognition and management oriented empirical and conceptual papers in leading journals and magazines. His present research focuses on statistical learning and its application to computer vision and image understanding and problem recognition

Trust challenges and issues of E-Government: E-Tax prospective

Dinara Berdykhanova

Asia Pacific University College of
Technology and Innovation
Technology Park Malaysia
Kuala Lumpur, Malaysia

Ali Dehghantanha

Asia Pacific University College of
Technology and Innovation
Technology Park Malaysia
Kualalumpur- Malaysia

Andy Seddon

Asia Pacific University College of
Technology and Innovation
Technology Park Malaysia
Kualalumpur- Malaysia

Abstract— this paper discusses trust issues and challenges have been encountered by e-government developers during the process of adoption of online public services. Despite of the apparent benefits as online services' immediacy and saving costs, the rate of adoption of e-government is globally below experts' expectations. A concern about e-government adoption is extended to trust issues which are inhibiting a citizen's acceptance of online public sector services or engagement with e-government initiatives. A citizen's decision to use online systems is influenced by their willingness to trust the environment and to the agency is involved. Trust makes citizens comfortable sharing personal information and making online government transaction. Therefore, trust is a significant notion that should be critically investigated in context of different E-Taxation models as part of E-Government initiatives. This research is proposing the implementation of Trusted Platform Module as a solution for achieving the high level of citizens' trust in e-taxation.

Keywords: E-Government, E-Taxation, Trust, Security, Trusted Platform Module.

I. INTRODUCTION

The phenomenon of the Internet has brought a transformational effect on the society. It has opened a new medium of communication for individuals and businesses and provided opportunities to communicate and get information in an entirely different way. The boosted usage of the Internet was initially due to private sector interests, but governments across the globe are now becoming part of this revolution. Governments worldwide have been making significant attempts to make their services and information available on the Internet [1].

The implementation of information technologies, particularly using Internet to improve the efficiency and effectiveness of internal government operations, communications with citizens, and transactions with both

individuals and organizations has brought new term –e-government or electronic government [2].

E-government can be defined as the use of primarily Internet-based information technology to enhance the accountability and performance of government activities. These activities include government's activities execution, especially services delivery, access to government information and processes; and citizens and organizations participation in the government [2].

Today E-Government offers a number of potential benefits to citizens. It gives citizens more control on how and when they interact with the government. Instead of visiting a department at a particular location or calling the government personnel at a particular time, citizens can choose to receive these services at the time and place of their choice [1]. As the result, various e-government initiatives have been taken with the objective to build services focused on citizens' needs and to provide more accessibility of government services to citizens [3]. In other words, the e-government can offer public service a truly standard, impersonal, efficient, and convenient manner for both service provider (the government) and service recipient (the citizens). In some cases a government agency can also be a service recipient of the e-government service.

In economic terms, the ability of citizens to access government services at anytime, anywhere helps to mitigate the transaction costs inherent in all types of government services [4]. In particular, on-line taxation is an important function of e-government since it is highly related to the life of citizens [5].

Electronic tax filing systems [6] are an e-government application which spreading rapidly all over the world. Those systems are particularly favorable for governments because they avoid many of the mistakes taxpayers make in manual filings, and they help to prevent tax evasion by data matching. The data warehouses developed using electronic tax filings allows tax inspectors to analyze declarations more thoroughly, and enable policy makers to develop fairer and more effective tax policies.

Due to factor that the taxes are crucial source of the budget revenue, the relationships between taxation and

technological developments have always been interactive, dynamic and complex. For the government taxation system is one of the most e-government applications where information technologies are highly penetrated [7].

The growing body of research recently presented that trust as essential element of successful e-government adoption process. It was found that lack of trust with respect to financial security and information qualities were among the barriers for a high level of adoption [9].

This paper is organized as follows. The next part will present understanding of "trust" in the e-government context, trust and e-taxation relationship and critical analysis of e-taxation models. The third part will propose solution addressing for the identified problems in e-taxation models.

II. LITERATURE REVIEW: TRUST AND E-TAXATION CONTEXT

Trust has been subject of researches in different areas, namely: technological, social, institutional, philosophical, behavioral, psychological, organizational, economic, e-commerce and managerial [10]. The literature also identifies trust as an essential element of a relationship when uncertainty, or risk, is present. Researchers are just beginning to empirically explore the role of trust in the e-government adoption [8]. Trust in the e-government is therefore composed of the traditional view of trust in a specific entity – government, as well as trust in the reliability of the enabling technology [11].

Despite the governments' growing investment in electronic services, citizens are still more likely to use traditional methods, e.g., phone calls or in-person visits, than the Web to interact with the government [11]. Therefore, the investigation of trust in the e-government is significant contribution in enabling cooperative behavior. Since the e-government applications and services more trustworthy, the higher level of citizens' engagement in public online services [1].

In e-taxation system trust is most crucial element of relationships, because during online transactions the citizens' vulnerable and private data are involved [12].

Internet tax filing was launched in Taiwan by the tax agency in 1998. Despite all the efforts aimed at developing better and easier electronic tax-filing systems, these tax-filing systems remained unnoticed by the public or were seriously underused in spite of their availability [12]. To analyze citizens' behavior [13] and [12] applied Technology Acceptance Model (TAM) as theoretical ground with the two major determinants: perceived usefulness and perceived ease of use of that system. In their research it been stated state that TAM's fundamental constructs do not fully reflect the specific influences of technological and usage-context factors that may alter the acceptance of the users. Moreover, their work has been stated that term as "trust" has a striking

influence on user willingness to engage in online exchanges of money and personal sensitive information.

Later, the new construct - "perceived credibility" has been proposed to add to TAM to enhance the understanding of an individual's acceptance behavior of electronic tax-filing systems [12].

The lack of "perceived credibility" [13] is manifested in people's concerns that the electronic tax-filing system will transfer their personal tax return information to the third parties without their knowledge or permission. Therefore, perceived fears of divulging personal information and users' feelings of insecurity provide unique challenges to find ways in which to develop users' perceived credibility of electronic tax-filing systems.

The following proposed components which help to measure of e-government success in Sweden and citizen satisfaction: e-government system quality, information quality, e-service quality, perceived usefulness, perceived ease of use, and citizen trust [3]. Moreover, it was stated citizens trust becomes one of the key components in enabling citizens to be willing to receive information and provide information back to the e-government system. In the government e-tax filing service, trust is defined as specific beliefs dealing with integrity, benevolence, competence, and predictability of government e-service delivery [3].

Trust is strongly associated with satisfaction with the e-government services, and satisfaction is related to citizens' perceptions about the service, such as the reliability of information provided by the government, the convenience of the service, etc. [14].

Trust is the expected outcome of e-government service delivery [15]. An absence of trust could be the reason for poor performance of e-government systems, and by improving service quality, trust can be restored. In other words, citizens must believe government agencies possess the astuteness and technical resources necessary to implement and secure these systems [8].

Users are concerned about the level of security present when providing sensitive information on-line and will perform transactions only when they develop a certain level of trust [3].

The link between security and trust has been studied in a number of researches. [10]. Therefore, citizens can trust in e-taxation system only when they perceive that their personal data are secured during online transactions. According to Brazil case study with the wide spreading of the e-taxation technology the question about security of online transactions was emerged. This item has been considered the "Achilles' heel" of the process, especially in the opinion of Tax Administrators at the developed countries [16].

The security issues have found in other countries as one of the main barrier to deep dissemination of public e-services of e-government. Security refers to the protection of information or systems from unsanctioned intrusions or

outflows. Fear of a lack of security is one of the factors that have been identified in most studies as affecting the growth and development of information systems [16]. Therefore, users' perception of the extent to which electronic tax-filing systems are capable of ensuring that transactions are conducted without any breach of security is an important consideration that might affect the use of the electronic tax-filing systems.

Observation of e-taxation system in USA [17] revealed significant obstacles to offering online tax services. The survey has shown that the percent of governments citing the issues of security which could reflect their interest in developing online transaction systems in USA.

Investigation of e-tax filing system in Japan by [18] has shown citizens' concerns about that national tax data contain sensitive personal and financial information. Any security breach will have negative impacts on the credibility of tax administration and public information privacy rights.

Taxpayers are very sensitive when they are filing their tax returns, since they need to provide a great deal of personal information. If they believe the tax authority is not opportunistic, then they will feel comfortable using this online service [18].

Trust in online services can be affected by new vulnerabilities and risks. While there are very few reliable statistics, all experts agree that direct and indirect costs of on-line crimes such as break-ins, defacing of web sites, spreading of viruses and Trojan horses, and denial of service attacks are substantial. Moreover, the impact of a concerted and deliberate attack on our digital society by highly motivated opponents is a serious concern [19].

In Australia for instance [20], launching of E-tax in project was successful, but not as expected. The efficiency issues are also provoked by other threads. Along with the massive growth in Internet commerce in Australia over the last ten years there has been a corresponding boom in Internet related crime, or cybercrime.

Despite of tremendous security programs and applications the number of reported security alerts does grow spectacularly and typically it increases by several per month [19].

As the Internet and underlying networked technology has continued to develop and grow then so has the opportunity for illicit behavior. Utilizing digital networks such as the Internet provides cyber criminals with a simplified, cost effective and repeatable mean to conduct rapid large scale attacks against global cyber community. Using methods such as email and websites eliminates the need for face-to face communication and provides the cyber criminal with a level of anonymity that reduces the perception of risk and also increases the appearance of legitimacy to a potential victim [21].

Therefore, today, more and more countries exploring e-government application to improve the access and delivery of government services to citizens are facing with

trust issues, whereas trust challenges impede further dissemination of public online service. On the other hand, citizens' trust is highly correlated with security of data while using e-taxation system applications. The lack of security of citizens' transactions can lead to refusal of interaction with e-government initiatives.

III. TRUSTED PLATFORM MODULE AS A TOOL TO RETAIN CITIZENS' TRUST

Having identified the importance of raising awareness and providing knowledge about security measurements to citizens, as a major factor in developing trust, this part of the paper will focus on new data security technologies and how those technologies can be used in a way which will enable citizens' trust while taking part in the e-government transactions.

February 2001 witnessed a major leap forward in the field of computer security with the publication of an innovative industry specification for "trusted platforms." This heralded a new era in significantly higher security for electronic commerce and electronic interaction than currently exists. What's the difference between a "platform" and a "trusted platform"? A platform is any computing device—a PC, server, mobile phone, or any appliance capable of computing and communicating electronically with other platforms. A Trusted Platform is one containing a hardware-based subsystem devoted to maintain trust and security between machines. This industry standard in trusted platforms is supported by a broad spectrum of companies including HP, Compaq, IBM, Microsoft, Intel, and many others. Together, they form the Trusted Computing Platform Alliance (TCPA) [22].

The TPM creates a hardware-based foundation of trust, enabling enterprises to implement, manage, and enforce a number of trusted cryptography, storage, integrity management, attestation and other information security capabilities Organizations in a number of vertical industries already successfully utilize the TPM to manage full-disk encryption, verify PC integrity, and safeguard data [23]. Moreover, TPM is bridging the gaps of the current solutions of data security Table 1 [24].

Table 1. Advantages of TPM over current security solutions

Threats	Current Solutions	Weaknesses	TPM Solutions
Data theft	Data encryption (EFS, VPN, encrypted email, etc.)	Encryption keys are stored on the hard disk and are susceptible to tampering	Protected storage of keys through hardware
Unauthorized access to platform	1. Username / Password 2. Biometrics and external tokens for user authentication	1. Subject to dictionary attacks 2. Biometrics can be spoofed 3. Authentication credentials not bound to platform	Protection of authentication credentials by binding them to platform
Unauthorized access to network	Windows network logon, IEEE 802.1x	1. Can be bypassed 2. Certificate can be spoofed 3. Authentication data is stored on the hard disk and is susceptible to tampering	1. PKI based method for platform authentication 2. Hardware protection of authentication data

The benefits from Table 1 are proving that TPM is reliable data security technology and implementing TPM in various e-government applications is suggested. Moreover, implementing TPM will provide robust and trustworthy security [19]. At the same time the embedding this technology does not requires significant investments. At the time of this writing, secure operating systems use different levels of hardware privilege to logically isolate programs and provide robust platform operation, including security functions.

As TPM was identified as a reliable technology for data security, the implementation of TPM in e-taxation system will provide built in protection of sensitive data. Since the citizens will experience robust security while using e-government initiatives, particularly e-taxation applications, the level of trust can significantly increase. The implementation of TPM will consist of the following consecutive stages:

- Framework development. It can include training for Tax agencies' workers manipulating with data been gained from e-tax websites. On this stage information about TPM on the websites should be uploaded, and then citizens would be aware that reliable security of online transactions would be conducted. Desired goals of implementation should be clarified, for instance the level of data security, increasing numbers of online service users, citizens' satisfaction and getting citizens trust.
- Test of TPM, can be provided while using TPM. The testing stage should include feedback from gathering Tax agency workers and service users.
- Evaluation of TPM. After testing the technology. On this stage feedback processing can represent whether the goals of TPM implementation are achieved or not.

IV. CONCLUSION

The delivery of information and services by the government online through the Internet or other digital means is referred to as e-Government. Governments all over the world have been making significant efforts in making their services and information available to the public through the Internet.

However, recent researches revealed that the success of e-Government efforts depends on not only technological excellence of services but other intangible factors. For instance, the term "trust" was frequently emerged in order to identify citizens' e-taxation services satisfaction. Analysis of different e-taxation models represents the trust as the key component of e-governments initiatives. Especially e-taxation, which requires sensitive and personal data while online transactions. Also, this research has shown that trust is occurring only when proper security is guaranteed.

As the solution to meet the trust requirements while using public online services TPM technology for data security was suggested. Low costs and security robustness make this approach more attractive security solution in online services compare to other existing security technologies. Implementation of TPM in e-taxation system will help to gain citizens trust and loyalty and will be conducted through several stages. The further direction of the research is evaluation of reliability of TPM technology and citizens satisfaction with the level of trust.

REFERENCES

- [1] Kumar Vinod, Bhasar Mukarji, Ifran Butt, "Factors for Successful e-Government Adoption: a Conceptual Framework", *Electronic Journal of e-Government*, Vol. 5, Issue 1, pp. 63 – 76, 2007. available online at www.ejeg.com. [Accessed November 1, 2009].
- [2] DeBenedictis, "E-government defined: an overview of the next big information technology challenge", *International Association for Computer Information Systems*, 2002.
- [3] Parmita Saha, "Government e-Service Delivery: Identification of Success Factors from Citizens' Perspective", Doctoral Thesis, *Lulea University of Technology*, 2008.
- [4] Kun Chan Lee, Melih Kirlog, Sangjae Lee, Gyoo Gun Lim, "User evaluations of tax filing web sites: A comparative study of South Korea and Turkey", *Online Information Review*, Volume 32 No. 6, pp. 842-859, 2008. Available at www.emeraldinsight.com. [Accessed October 30, 2009].
- [5] Ing-Long Wu, Chen Jian-Liang, "An extension of Trust and TAM model with TPB in the initial adoption of on-line tax: An empirical study", *International Journal of Human-Computer Studies*, Vol. 62, pp. 784–808, 2005.
- [6] T.S. Manly, D.W Thomas and C.M. Ritsema, "Attracting nonfilers through amnesty programs: internal versus external motivation", *Journal of the American Taxation Association*, Vol. 27, pp. 75-95, 2005.
- [7] www.americatrading.ca, eTaxation, [Accessed October 4, 2009].
- [8] France Belanger, Lemuria Carter, "Trust and risk in the e-government adoption", *Journal of Strategic Information Systems*, Issue 17, pp. 165–176, 2008.
- [9] Helle Zinner Henriksen, "Fad or Investment in the Future: An Analysis of the Demand of e-Services in Danish Municipalities", *The Electronic Journal of e-Government*, Vol. 4 Issue 1, pp 19 - 26, 2006. Available online at www.ejeg.com. [Accessed November 1, 2009].
- [10] Rana Tassabehji, Tony Elliman, "Generating Citizen Trust in the e-government using a Trust Verification Agent", *European and Mediterranean Conference on Information Systems (EMCIS)*, July 6-7, 2006, Costa Blanca, Alicante, Spain.
- [11] Lemuria Carter, Belanger France, "The utilization of e-government services: citizen trust, innovation and acceptance factors", *Information Systems Journal*, Vol. 15, Issue 1, pp. 5–25, 2005.
- [12] Jen-Ruei Fu, Cheng Kiang Farn, Wan Pin Chao, "Acceptance of electronic tax filing: A study of taxpayer intentions", *Information & Management*, Vol. 43, pp. 109–126, 2006.
- [13] Yei-Shung Wang, "The adoption of electronic tax filing systems: an empirical study", *Government Information Quarterly*, Volume 20, pp. 333–352, 2002.
- [14] Hisham Alsaghier, Marilyn Ford, Anne Nguyen, Rene Hexel "Conceptualising Citizen's Trust in e-Government: Application of Q Methodology", *Electronic Journal of e-Government*, Vol. 7, Issue 4, pp.295-310, 2009. Available online at www.ejeg.com. [Accessed November 1, 2009].

- [15] Eric W. Welch, "Linking Citizen Satisfaction with e-Government and Trust in Government", *Journal of Public Administration Research and Theory*, Volume 15, Issue 3, 2004.
- [16] Maria Virginia de Vasconcellos, Maria das Graças Rua, "Impacts of Internet use on Public Administration: A Case Study of the Brazilian Tax Administration", *The Electronic Journal of e-Government*, Vol. 3, Issue 1, pp 49-58, 2005. Available online at www.ejeg.com. [Accessed November 2, 2009].
- [17] Bruce Rocheleau and Liangfu Wu, "E-Government and Financial Transactions: Potential versus Reality", *The Electronic Government Journal*, Vol. 2, Issue 4, pp. 219-230, 2005. Available online at www.ejeg.com. [Accessed November 2, 2009].
- [18] Akemi Takeoka Chatfield, "Public Service Reform through e-Government: a Case Study of 'e-Tax' in Japan", *Electronic Journal of e-Government*, Vol.7 Issue 2, pp.135 – 146, 2009. Available online at www.ejeg.com. [Accessed November 2, 2009].
- [19] Boris Balacheff, Liquan Chen, Siani Pearson, David Plaquin, Graeme Proudler, "Trusted Computing Platforms: TCPA Technology in Context", Prentice Hall PTR, 2002.
- [20] Mehdi Khourow-Pour, "Cases of electronic commerce technologies and applications", Idea Group Publishing, 2006.
- [21] P. Hunton, "The growing phenomenon of crime and the internet: A cybercrime execution and analysis model", *Computer law and security review*, Issue 25, pp. 528-535, 2009.
- [22] Sean Smith, "Trusted Computing Platforms: Design and Application", Dartmouth College, 2005.
- [23] Trusted computer Group, "Enterprise Security: Putting the TPM to Work", 2008. Available online at www.trustedcomputinggroup.org. [Accessed October 20, 2009].
- [24] Sundeep Bajkar, "Trusted Platform Module (TPM) based Security on Notebook PCs", Mobile Platforms Group Intel Corporation, 2002.

Machine Learning Approach for Object Detection - A Survey Approach

N.V. Balaji

Department of Computer Science, Karpagam
University,
Coimbatore, India.

Dr. M. Punithavalli

Department of Computer Science, Sri Ramakrishna
Arts College for Women,
Coimbatore, India.

Abstract---Object detection is a computer technology related to computer vision and image processing to determine whether or not the specified object is present, and, if present, determine the locations and sizes of each object. Depending on the machine learning algorithms, we can divide all object detection methods as Generative Methods and Discriminative Methods. The concept of object detection is being an active area of research and it is rapidly emerging since it is used in many areas of computer vision, including image retrieval and video surveillance. This paper presents a general survey which reviews the various techniques for object detection and brings out the main outline of object detection. The concepts of image detection are discussed in detail along with examples and description. The most common & significant algorithms for object detection are further discussed. In this work an overview of the existing methodologies and proposed techniques for object detection with future ideas for the enhancement are discussed.

Keywords---Object Detection, Support Vector Machine, Neural Networks, Machine Learning.

I. INTRODUCTION

Extracting a feature vector of a given object and object detection using the feature vector using pattern matching technique is the main goal for object detection [2]. Object detection is to determine whether or not the object is present, and, if present, determine the locations and sizes of each object.

The most common approaches are image feature extraction, feature transformation and machine learning where image feature extraction is to extract information about objects from raw images.

Classification of patterns, object identification and its description, are important tribulations to be concentrated upon in a variety of engineering and scientific disciplines such as biology, psychology, medicine, marketing, computer vision, artificial intelligence, and other remote sensing. Watanabe [1] defines a pattern as opposite of a chaos, that is, it is an entity, vaguely defined and that could be given a name. For instance, a pattern could be a fingerprint image, a handwritten cursive word, a human face, or a speech signal. Given a pattern, the object detection may consist of one of the following two tasks [2] either the supervised classification in which the input pattern is identified as a member of a predefined class or the unsupervised classification, which the pattern is assigned to a previously unknown class.

The recognition problem is being posed as a classification task, where the classes are either defined by the system designer or are learned based on the similarity of patterns. Interest in the area of object detection has been renewed

recently due to emerging applications which are not only challenging but also computationally more demanding. These applications include data mining, document classification, financial forecasting, organizing and retrieval of multimedia databases, and biometrics and also the other fields where the need of the image detection is high.

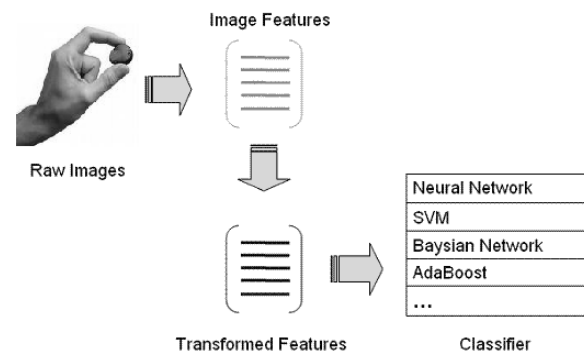


Figure 1. Description for the Image Detection

The recognition problem is being posed as a classification task, where the classes are either defined by the system designer or are learned based on the similarity of patterns. Interest in the area of object detection has been renewed recently due to emerging applications which are not only challenging but also computationally more demanding. These applications include data mining, document classification, financial forecasting, organizing and retrieval of multimedia databases, and biometrics and also the other fields where the need of the image detection is high.

II. LITERATURE SURVEY

Extraction of a reliable feature and improvement of the classification accuracy have been among the main tasks in digital image processing. Finding the minimum number of feature vectors, which represent observations with reduced dimensionality without sacrificing the discriminating power of pattern classes, along with finding specific feature vectors, has been one of the most important problems in the field of pattern analysis.

In the last few years, the problem of recognizing object classes received growing attention in both variants of whole image classification and object localization. The majority of existing methods use local image patches as basic features [3]. Although these work well for some object classes such as motor-bikes and cars, other classes are defined by their shape and therefore better represented by contour features.

In many real world applications such as pattern recognition, data mining, and time-series prediction, we often confront difficult situations where a complete set of training sample is not given when constructing a system. In face recognition, for example, since human faces have large variations due to expressions, lighting conditions, makeup, hairstyles, and so forth, it is hard to consider all variations of face in advance.

In many cases, training samples are provided only when a system misclassifies objects; hence the system is learned online to improve the classification performance. This type of learning is called incremental learning or continuous learning, and it has recently received a great attention in many practical applications.

In pattern recognition and data mining, input data often have a large set of attribute. Hence, the informative input variables (features) are first extracted before the classification is carried out. This means that when constructing an adaptive classification system, we should consider two types of incremental learning: one is the incremental feature extraction, and the other is incremental learning classifiers.

A. A hybrid object detection technique

As discussed by M. Paul et. al., in [9] the adaptive background modeling based object detection techniques are widely used in machine vision applications for handling the challenges of real-world multimodal background. But they are forced to detailed environment due to relying on environment precise parameters, and their performances also alter across dissimilar operating speeds. The basic background calculation is not appropriate for real applications due to manual background initialization prerequisite and its incapability to switch cyclical multimodal background. It shows better firmness across different operating speeds and can better abolish noise, shadow, and trailing effect than adaptive techniques as no model adaptability or environment related parameters are involved. The hybrids object detection technique for incorporating the strengths of both approaches. In this technique, Gaussian mixture models is used for maintaining an adaptive background model and both probabilistic and basic subtraction decisions are utilized for scheming reasonably priced neighbor hood statistics for guiding the final object detection decision.

B. Moving Object Detection Algorithm

Zhan Chaohui et. al., projected in [10], the first point in moving object detection algorithm is the block-based motion assessment is used to attain the common motion vectors, the vectors for every block, where the central pixel of the block is considered as the enter crucial point. These motion vectors are used to sense the border line blocks, which contain the border of the object. Presently on, the linear exclamation is used to make the coarse motion field an impenetrable motion field, by this way to eliminate the chunk artifacts. This possession can also be used to sense whether the motion field is uninterrupted or not. This sophisticated impenetrable motion field is used to define detail limitations in each boundary block. Thus the moving object is detected and coded.

C. Restricted Bayesian networks

This approach presented by Schneiderman et. al., in [4, 5, 6 and 7] attempts to study the structure of a Bayesian network

to carry out the generative task. The problem of sophistication the construction of a Bayesian network is known to be NP-Hard, and therefore it is restricted to the structure of the concluding network to a known form of arrangement to gain tractability.

The initial phase is to mine feature information from the object. Schneiderman has discussed by using three level wavelet transform to convert the contribution image to spatial occurrence in sequence. One then constructs a set of histograms in both position and intensity. The concentration values of each wavelet layer need be quantized to fit into an inadequate number of bins. One difficulty encountered in the premature execution of this method was the lacking of high power regularity information in the objects. With a linear quantization scheme the higher energy bins had primarily singleton values, this leads to a problem when a prior is introduced to the bin, as the actual count values are lost in the introduced prior. To extract this exponential quantization technique was employed to spread the power evenly between all the bin levels.

D. Cluster-Based Object Detection

The cluster based object detection was proposed by Rikert, Jones, and Viola [8]. In this methodology, the information about the object is learned and used for classification. The objects are transformed and then build a mixture of Gaussian model. The transformation is done based on the result of k-means clustering applied to the transformed object. In the initial pace the object is distorted using a multi-directional steerable pyramid. The result of the pyramid is then compiled into a succession of quality vectors self-possessed of the foremost coat deposit pixel, and the pixels from higher in pyramid resized without interruption. For reasonably sized patches this quickly becomes intractable.

E. Rapid Object Detection using a Boosted Cascade of Simple Features

Paul Viola et. al., describe in [11], as a machine learning approach for object detection which is capable of processing images tremendously rapid and achieving high detection rates. This work is illustrious by three key contributions. The initial is the prologue of an original object representation called the integral object which allows the features used by the detector to be computed very quickly. The author developed a learning algorithm, based on Ada Boost, which selects a small number of critical visual features from a superior set and yields enormously efficient classifiers [12]. The third contribution is a method for combining increasingly more complex classifiers in a "cascade" which allows background regions of the object to be quickly discarded while spending more calculation on showing potential object-like regions. The flow can be viewed as an object specific focus of concentration mechanism which dissimilar to preceding approaches that provides statistical guarantees that superfluous regions are improbable to contain the object of interest.

F. Template Matching Methods

Huang T.S et. al., described the template matching methods that uses standard patterns of objects and the object parts to portray the object globally or as diverse parts. Correlations get struck between the input image and patterns subsequently computed for detection. Gavrilu [16] proposed an object detection scheme that segments forefront regions

and extracts the boundary. Then the algorithm searches for objects in the image by matching object features to a database of templates. The matching is realized by computing the average Chamfer detachment amid the template and the edge map of the target image area. Wren et al. [18] described detailed on a top-down person detector based on template-matching. However, this approach requires field specific scene analysis.

G. Object Detection Using Hierarchical MRF and MAP Estimation

Qian R.J et. al., projected this method in [15] which presents a new scale, position and direction invariant approach to object detection. The technique initially chooses concentration on regions in an object based on the region detection consequence on the object. Within the attention regions, the method then detects targets that combine template matching methods with feature-based methods via hierarchical MRF and MAP estimation. Hierarchical MRF and MAP inference supply a stretchy framework to integrate various visual clues. The amalgamation of template corresponding and feature detection helps to accomplish robustness against multifaceted backgrounds and fractional occlusions in object detection.

H. Object Detection and Localization using Local and Global Features

The work proposed by Kevin Murphy et. al., in [21] describes more advanced method of object detection and localization using local and global features of an image. Traditional approaches to object detection only look at the local pieces of the image, whether it can be within a sliding window or the regions around an interest point detector. When this object of interest is small or the imaging conditions are otherwise unfavorable, such local pieces of the image can become indistinct. This ambiguity can be reduced by using global features of the image – which we call as a “gist” of the scene. The object detection rates can be significantly improved by combining the local and global features of the image. This method also results in large increase of speed as well since the gist is much cheaper to compute than the local detectors.

I. Object Detection from HS/MS and Multi-Platform Remote Sensing Imagery

Bo Wu et.al, put forth a technique in [22] that integrates biologically and geometrically inspired approaches to detect objects from hyperspectral and/or multispectral (HS/MS), multiscale, multiplatform imagery. First, dimensionality reduction methods are studied and implemented for hyperspectral dimensionality reduction. Then, a biologically stimulated method S-LEGION (Spatial-Locally Excitatory Globally Inhibitory Oscillator Network) is developed for object detection on the multispectral and dimension reduced hyperspectral data. This method provides rough object shapes. Geometrically inspired method, GAC (Geometric Active Contour), is employed for refining object boundary detection on the high resolution imagery based upon the initial object shapes provided by S-LEGION.

J. Binary Partition Tree for Object Detection

This proposal suggested by V. Vilaplana et. al., in [23], discusses the use of Binary Partition Tree (BPT) for object detection. BPTs are hierarchal region based representation of

images. They define a reduced set of regions that covers the image support and that spans various levels of resolution. They are attractive for object detection as they enormously reduce the search space. In [23], several issues allied to the use of BPT for object detection are examined. Analysis in the compromise between computational complexity reduction and accuracy in accordance with the construction of binary tree lead us to define two parts in BPT: one providing the accuracy and the other representing the search space for the task of object detection. In turn, it includes an analysis and objectively compares various similarity measures for the tree construction. This different similarity criterion should be used for the part providing accuracy in the BPT and for the part defining the search space. Binary Partition Tree concentrates in a compact and structured representation of meaningful regions that can be extracted from an image. They offer a multi-scale representation of the image and define the translation invariant 2-connectivity rule among regions.

K. Statistical Object Detection Using Local Regression Kernels

This novel approach was proposed by Hae Jong Seo and Peyman Milanfar in [24] to the problem of detection of visual similarity between a template image and patches in the given image. The method is based on the computation of the local kernel of the template, which measures the likeness of a pixel to its surroundings. This kernel is then used as a descriptor from which features are extracted and compared against analogous features from the target image. Comparison of the features extracted is carried out using canonical correlations analysis. The overall algorithm yields a scalar resemblance map (RM). This resemblance map indicates the statistical likelihood of similarity between a given template and all target patches in an image. Similar objects with high accuracy can be obtained by performing statistical analysis on the resulting resemblance map. This proposed method is robust to various challenging conditions such as partial occlusions and illumination change.

L. Spatial Histogram based Object Detection

Hongming Zhang et. al., describes in [25], that feature extraction plays a major role for object representation in an Automatic object detection system. The spatial histogram preserves the object texture and shape simultaneously as it contains marginal distribution of image over local patches. In [25], methods of learning informative features for spatial histogram-based object detection were proposed. Fisher criterion is employed to measure the discriminability of the spatial histogram feature and calculates features correlations using mutual information. An informative selection algorithm was proposed in order to construct compact feature sets for efficient classification. This informative selection algorithm selects the uncorrelated and discriminative spatial histogram features and this proposed method is efficient in object detection.

M. Recursive Neural Networks for Object Detection

M. Bianchini et. al., put forth an algorithm in [26], a new recursive neural network model for object detection. This algorithm is capable of processing directed acyclic graphs with labeled edges, which address the problem of object detection. The preliminary step in an object detection system is the detection. The proposed method describes a graph-based

representation of images that combines both spatial and visual features. The adjacency relationship between two homogeneous regions after segmentation can be determined by the edge between the two nodes. This edge label collects information on their relative positions, whereas node labels contain visual and geometric information on each region (area, color, texture, etc). These graphs are then processed by the recursive model in order to determine the eventual presence and the position of objects inside the image. The proposed system is general and can be employed to any object detection systems, since it does not involve any prior knowledge on any particular problem.

N. Object Detection Using a Shape Codebook

Object detection by Xiaodong Yu et. al., in [27], presents a method for detecting categories of object in real world images. The ultimate aim is to localize and recognize instances in the training images of an object category. The main contribution of this work is a novel structure of the shape code-book for object detection. The code book entry consists of two components: a shape codeword and a group of associated vectors that specify the object centroids. The shape codeword is such that it can be easily extracted from most image object categories. A geometrical relationship between the shape codeword is stored by the associated vectors. The characteristics of a particular object category can be specified by the geometrical relationship.

Triple-Adjacent-Segments (TAS), extracted from image edges is used as a shape codeword. Object detection is carried out in a probabilistic voting framework. This proposed method has drastically lower complexity and requires noticeably less supervision in training.

O. Contour-based Object Detection in Range Images

This approach investigated by Stefan Stiene et. al., in [28], projects a novel object recognition approach based on range images. Due to the insensitivity to illumination, range data is most suited for reliable outline extraction. This determines (silhouette or contour descriptions) as good sources for object recognition. Based on a 3D laser scanner, contour extraction performed using floor interpretation. Feature extraction is done using a new fast Eigen-CSS method and a supervised learning algorithm. It proposes a complete object recognition system. This recognition system was found to be tested successful on range images captured with the help of mobile robot. This results is compared with standard techniques i.e., Geometric features, Border signature method, and the angular radial transformation. The Eigen-CSS method is found to be more efficient than the best one by an order of magnitude in feature extraction time.

III. FUTURE ENHANCEMENT

The object detection methodologies are improving day by day as its need is hastily growing. Hence the techniques consider feature extraction where the Principal Component Analysis and Linear Discriminant Analysis are the most common approaches available. In object detection systems, the complete set of samples is not given at the time constructing the system. Instead, more and more samples are added whenever the system misclassifies objects and hence the system is learned online to improve the classification performance. This type of learning is called incremental

learning or continuous learning. This problem is to introduce Incremental Linear Discriminant Analysis as the feature extraction technique to object detection and hence improving the classification performance to the great height.

The overall outcome of the proposed work is to implement a variation in the existing feature extraction system LDA and to develop a new system ILDA which increases the classification performance to a great height. Also the system should take new samples as input online and learn them quickly. As a result of this incremental learning process, the system will have a large set of samples learned and hence will decrease the chance of misclassifying an object.

IV. CONCLUSION

This paper attempts to provide a comprehensive survey of research on object detection and to provide some structural categories for the methods described. When appropriately considered, it is been reported that, on the relative performance of methods in so doing, it needs the awareness that there is a lack of uniformity in how methods are evaluated and so it is reckless to overtly state that which methods indeed have the lowest error rates. As a substitution, it can be urged to the members of the community to expand and contribute to test sets and to report results on already available test sets. The community needs to more seriously considered for systematic performance evaluation. This would allow users and the researchers of the object detection algorithms to identify which ones are aggressive in which particular domain. It will also prompt researchers to produce truly more effective object detection algorithms.

REFERENCES

- [1] S. Watanabe, Pattern Recognition: Human and Mechanical. New York: Wiley, 2005.
- [2] Ilin, R. Kozma, R. Werbos, P.J, "Beyond Feedforward Models Trained by Backpropagation: A Practical Training Tool for a More Efficient Universal Approximator", IEEE Transactions on Neural Networks, Vol. 19, No. 6, June 2008.
- [3] Culp, M. Michailidis, G, "Graph-Based Semisupervised Learning ", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 30 Issue: 1, 2008.
- [4] H. Schneiderman, "A Statistical Approach to 3d Object Detection Applied to Faces And Cars", 2000.
- [5] H. Schneiderman. Learning statistical structure for object detection, 2003.
- [6] H. Schneiderman. Feature-centric evaluation for efficient cascaded object detection, 2004.
- [7] H. Schneider man. Learning a restricted Bayesian network for object detection. 2004.
- [8] P. Viola T. Rikert and M. Jones. A cluster-based statistical model for object detection.1999.
- [9] Haque, M. Murshed, M. Paul, M, "A hybrid object detection technique from dynamic background using Gaussian mixture models", Gippsland Sch. of Inf. Technol., Monash Univ., Clayton, VIC, Multimedia Signal Processing, 2008 IEEE 10th Workshop, Oct. 2008.
- [10] Zhan Chaohui, Duan Xiaohui, Xu Shuoyu, Song Zheng and Luo Min, "An Improved Moving Object Detection Algorithm Based on Frame Difference and Edge Detection", ICIG 2007. Fourth International Conference, Aug. 2007.
- [11] Paul Viola and Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE 2001.
- [12] Yoav Freund and Robert E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", In Computational Learning Theory: Eurocolt '05, Springer-Verlag, 2005.
- [13] Sharaf, R. Noureldin, A, "Sensor Integration for Satellite-Based Vehicular Navigation Using Neural Networks", IEEE Transactions on Neural Networks, March 2007.

- [14] Juyang Weng Yilu Zhang Wey-Shiuan Hwang , "Candid covariance-free incremental principal component analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003.
- [15] Qian R.J and Huang T.S, "Object detection using hierarchical MRF and MAP estimation", Vision and Pattern Recognition, 2007. Proceedings. 2007 IEEE Computer Society Conference.
- [16] D.M. Gavrila and V. Philomin, "Real-time object detection for smart vehicles", IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR), pp. 87-93, 1999.
- [17] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 511-518, 2001.
- [18] C. R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "PFinder: real-time tracking of the human body", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 780-785, 1997.
- [19] M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer, "Theoretical foundations of the potential function method in pattern recognition learning," Automat. Remote Contr., vol. 25, pp. 917-936, 2004.
- [20] L. I. Rozonoer, "The probability problem of pattern recognition learning and the method of potential functions," Automat. Remote Contr., vol. 25, 2004.
- [21] Kevin Murphy, Antonio Torralba, Daniel Eaton, and William Freeman. Object detection and localization using local and global features. 2007.
- [22] Bo Wu, Yuan Zhou, Lin Yan, Jiangye Yuan, Ron Li, and DeLiang Wang, "Object detection from HS/MS and multi-platform remote sensing imagery by the integration of biologically and geometrically inspired approaches," ASPRS 2009 Annual Conference, Baltimore, Maryland. March 9-13, 2009.
- [23] V. Vilaplana, F. Marques and P. Salembier, "Binary Partition Tree for object detection", Image Processing, IEEE Transactions on Image Processing, vol. 17, no. 11, pp. 2201-2216, 2008.
- [24] Hae Jong Seo and Peyman Milanfar, "Using local regression kernels for statistical object detection," IEEE Transaction on Pattern Analysis and Machine Intelligence 2008.
- [25] Hongming Zhang, Wen Gao, Xilin Chen, and Debin Zhao, "Learning informative features for spatial histogram-based object detection," IEEE 2005. International Joint Conference on Neural Networks, Montreal, Canada, July 31-August 04, 2005.
- [26] M. Bianchini, M. Maggini, L. Sarti, and F. Scarselli, "Recursive neural networks for object detection," IEEE Transaction on Pattern Analysis and Machine Intelligence 2004.
- [27] Xiaodong Yu, Li Yi, Cornelia Fermuller, and David Doermann, "Object detection using a shape code-book," IEEE, 2007.
- [28] Stefan Stiene, Kai Lingemann, Andreas Nuchter, and Joachim Hertzberg, "Contour-based Object detection in range images," IEEE 2006.

AUTHORS PROFILE



N.V. Balaji has obtained his Bachelor of Science in Computer Science from Sri Ramasamy Naidu Memorial College, Sattur in 1997 and Master of Science in Computer Science from Dr. GRD College of Science in 1997. Now he is doing Ph.D., in Bharathiar University. He commences more than nine years of experience in teaching field moreover industrial experience in Cicada Solutions, Bangalore. At present he is working as Asst.

Professor & Training Officer at Karpagam University. His research interests are in the area of Image Processing and Networks. He presented number of papers in reputed national and international journals and conferences.



Dr. M. Punithavalli received the Ph.D degree in Computer Science from Alagappa University, Karaikudi in May 2007. She is currently serving as the Adjunct Professor in Computer Application Department, Sri Ramakrishna Engineering College, Coimbatore. Her research interest lies in the area of Data mining, Genetic Algorithms and Image Processing. She has published more than 10 Technical papers in International, National

Journals and conferences. She is Board of studies member various universities and colleges. She is also reviewer in International Journals. She has given many guest lecturers and acted as chairperson in conference. Currently 10 students are doing Ph.D., under her supervision.

Performance comparison of SONET, OBS on the basis of Network Throughput and Protection in Metropolitan Networks

Mr. Bhupesh Bhatia
Assistant Professor
Northern India Engineering College,
New Delhi, India

R.K. Singh
Officer on special duty,
Uttarakhand Technical University,
Dehradun (Uttarakhand), India.

Abstract— In this paper we explore the performance of SONET/SDH & OBS architectures connected as mesh topology, for optical metropolitan networks. The OBS framework has been widely studied in past days because it achieves high traffic throughput & high resource utilization. A brief comparison between OBS & SONET is studied. The results are based on analysis of simulations and we present a comparison between OBS architectures (with centralized & distributed scheduling schemes), SONET & NG-SONET.

Keywords— Add Drop Multiplexers; LCAS latency; Over Provisioning; WR-OBS; JET-OBS; Network Protection.

I. INTRODUCTION

SONET & SDH are multiplexing protocols which are used to send the digital bits over the fiber optics cable with the help of LASER or LED. If the data rates could be compensated in terms of speed then it could be transmitted via electrical interface. These are designed for the replacement of PDH systems used for telephonic data and other data over the same fiber optics cable at an improved speed. SONET allowed the user to communicate with different user's at different speeds i.e. in the asynchronous speed. So it is not just as the communication protocol but also a transport protocol. So it becomes the first choice to work in the asynchronous transfer mode. So they are used widely in the world. The SONET is used in the United States and Canada and SDH is used in the rest of the world. [5]

OBS is a kind of switching which lies in between the optical circuit and optical packet switching. This type of switching is appropriate for the provision of light paths from one to another node for many services/clients. It operates at the sub level wavelength and it is designed to improve the utilization of wavelength by quick setup. In this the data from the client side is aggregated at the network node and the sends on the basis of assembly/aggregation algorithm. [5]

II. OPTICAL PACKET SWITCHING NETWORK AND TOPOLOGIES

SONET network architecture made up of 2x2 optical network nodes, interconnected uni-directionally. They have optical add-drop Multiplexers. The higher node allows user to

connect to other sub-networks by Wavelength Division Multiplexing. The switching is controlled by electronic logic circuits which are based on packet-by-packet, which is determined only by header processing. [1]

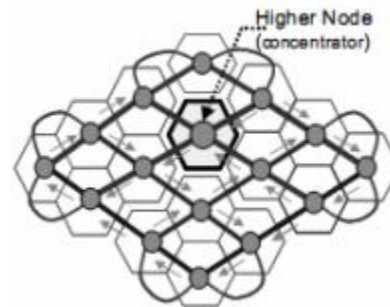


Figure 1. Unidirectional Mesh topology optical network

The overall switching time is less than two microseconds for every packet and is independent of payload size. This architecture helps to use the deflection routing to avoid collisions and there is no need for further buffering and thus cost reduces [2][3].

This provides the optical nodes to be operated asynchronously. Our solution is given for MAN access and distribution, having 15km length and networks having less than 48 nodes [2].

Mesh topology is selected for the analysis of throughput and to find the load on each node. The motive is to find which links are more frequently used and should be secured to avoid loss of critical service. These considerations also include the cost parameter.

III. BASIC THEORY AND PARAMETERS

The total capacity that can be given for a network is shown by (1) where, H is the total average number of hops from origin to destination, N number of nodes and S link capacity; the factor 2 is used because each node has possibility of two outputs [2][3].

$$C_t = \frac{2.N.S}{H} \quad (1)$$

If we consider Poisson distribution eq., every node generates the uniform traffic to each node and the link is of unidirectional nature. The no. of users in this network is $N(N-1)$. So the capacity can be given by :-

$$C_u = \frac{2.S}{H.(N-1)} \quad (2)$$

If there is any link failure, the network capacity decreases and if total links of $2N$ and m links are failed, then the capacity can be given as :-

$$C_t = \frac{(2N-m).S}{H} \quad (3)$$

If the network load seems to be L_c and the capacity be C_t then the network throughput can be given as :-[4]

$$T_p = C_t.L_c. \quad (4)$$

To determine the throughput for each destination node and then take an average, a general expression for T_p can be written as [6] :-

$$T_p = \frac{\sum_{i=1}^N T_{pi}}{N} \quad (5)$$

where i = destination node

T_{pi} = partial throughput to that node

N = total number of nodes

IV. SIMULATION METHODS AND NETWORK CONFIGURATIONS

Here we choose the mesh topology MSq-24, MS-32, MSq-48, with 24, 32, 48 nodes with bit-rate is 4.2 Gb/s, and link length of 15km.

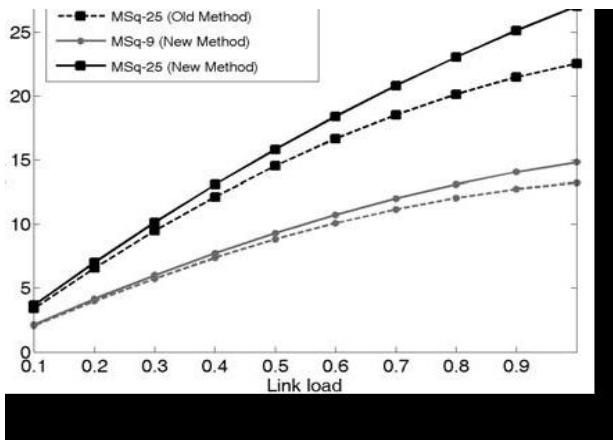


Figure 2. Comparative Throughput for mesh networks using old and new methods

Here it is supposed that each node is generating the equal traffic to every other node. Applications can be defined as the

total number of packets transmitted from a node to all other Connected nodes, and the sum of all applications is the total traffic load on the network. For the analysis of the protection, we take only single link failure. The SONET network traffic graphs were obtained using the Network Simulator software. [6][7][8]

V. RESULTS AND DISCUSSION

The throughput for mesh topology is shown in the figure.3. Here, we can observe that SONET performed well in the mesh network and brilliant in the condition of higher number of nodes. From this we can conclude that mesh topology is providing the high capacity without considering the cost of installation. We can see the traffic analysis of MS-24, MS-32, MS-48 and the protocols used in this analysis is “store and forward”.

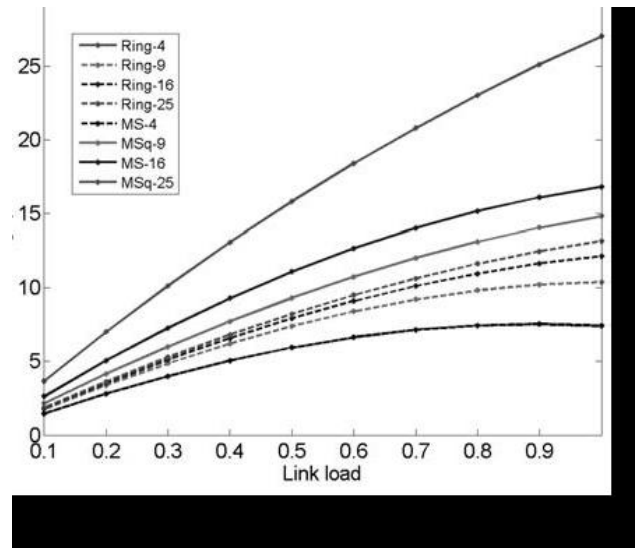


Figure 3. Comparative Throughput for mesh using the new method

Although in the above mentioned technique i.e. Store & Forward, the sent packets have to wait so as to provide them a shortest path for their destination, it doesn't matter because here we are just considering the utilization of links and their corresponding distribution of traffic. But ideally we should restrict ourselves to overload the certain links so as to minimize the failures, and we must take decision that where to apply protection mechanisms.

VI. NETWORK PROTECTION AND FAILURE ANALYSIS

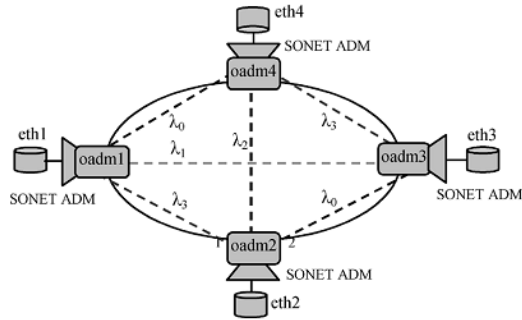
In mesh network, the links which are failed and less used, made a slight change in the performance of the network. The simulations include the MSq-24, MS-32, and MSq-48. We observe that in mesh topology the performance and the throughput reduced but the rate of reduction is almost half as compare to ring topology. In the mesh topology some more features are seen like protection of network, location of failure and finally restoration. So all such problems are reduced in the mesh topology.

VII. NG-SONET (NEXT GENERATION SONET)

NG-SONET is another approach which is most recent and in this there is provision of the carriers for optimizing the

allocation of bandwidth and uses the unused and fragmented capacity in the SONET ring. It also matches the better client rate. It uses some new protocols to accomplish these tasks such as generic framing for encapsulating data and virtual catenation for using fragmented bandwidth and (LACS) link capacity adjustment for resizing of existing links [9][10]. But it has some drawbacks which are:-

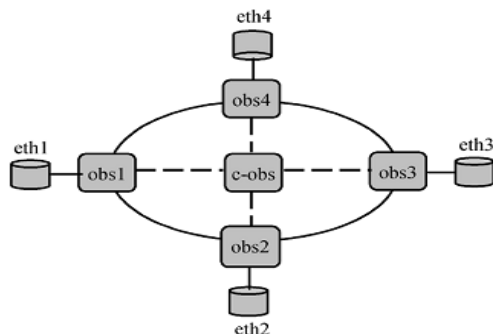
1. Over provisioning of links in case of Ethernet usage.
2. LCAS latency.



SONET & NG-SONET Network Models [14]

VIII. WR-OBS (WAVELENGTH ROUTED OBS)

In WR-OBS, the control packets are processed at a central node to determine the actual path to send the packets at the destination. The acknowledgements are sent to the source nodes and decided whether these are destroyed or transmit the data bursts. So this technique is best for optimal path selection which in turn gives the congestion control and helps in balancing the traffic over links. It has time delay consists of aggregation time and connection establishment time. It provides less delay than SONET & NG-SONET for low bandwidth links. This is due to the Ethernet packet transmissions are independent of time slot and frames. [11][12][13]



OBS-JET & WR-OBS Network Models [14]

Offset time delay $T_o = 3t_p$

t_p = processing time at each hop

IX. JET-OBS (JUST ENOUGH TIME)

In this an offset time is transmitted before the data burst is sent and processed electronically at each node for preserving the resources for the each data bursts. But the offset time must

be carefully chosen so that there should not be problem aroused of queuing and delay problem between the hops [11][12][13].

It has two types of delays:-

1. Aggregation Delay T_i
2. Offset time delay T_o

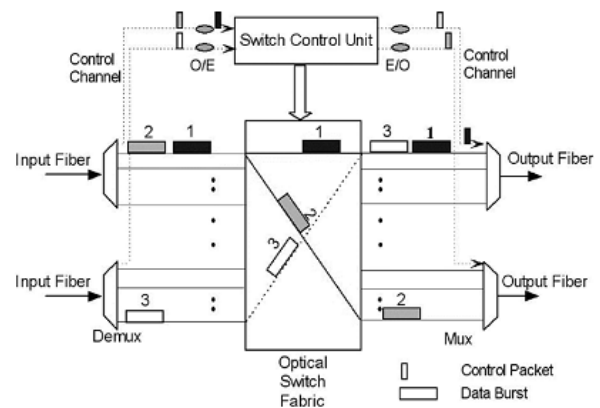
Where $T_i = N / \lambda - i / \lambda$

N = average number of packets

λ = mean arrival of packets

and $T_o = 3t_p$

t_p = processing time at each hop



Architecture of OBS-JET core node [14]

X. COMPARISON

OBS is a kind of switching which lies in between the optical circuit and optical packet switching whereas SONET is multiplexing protocols which are used to send the digital bits over the fiber optics cable [5]. OBS has three wavelengths for data and one wavelength for control channel whereas SONET has all four wavelengths available for data transmissions. OBS has data loss due to scheduling contentions while in SONET data loss is due to excessive delays [15]. OBS is of two types Just Enough Time (JET) OBS & Wavelength Routed (WR) OBS while SONET is of one type NG-SONET. OBS is not good for ring model network while SONET works best in ring network. OBS uses deflection routing to avoid contention whereas in SONET there is no such algorithm. OBS uses the forwarding tables for mapping the bursts whereas SONET has no such facility. OBS is preferred for busty traffic whereas SONET is not preferred for a busty traffic [15].

XI. CONCLUSION

We have studied and analyzed the capacity and throughput of SONET & OBS in mesh topology and have reached at the decision that mesh topology is better than the ring topology. If we talk about the protection, then we observe that the failure of links has more impact on ring topology instead of mesh topology. Also in the mesh topology, the impact on capacity due to failed links is much less and is less critical than the ring topology and this confirm that the mesh topology is robust in

nature. Also other features such as protection, restoration, and location of fault technique are absent in ring topology.

XII. FUTURE WORK

For the future prospects, the OBS will be studied and performance will be observed on the different networks like hybrid networks and other kind of topologies. Also their throughput and capacity will also be studied and if found to be satisfactory then the above study will be improved or may be replaced. Along with it, edge delay analysis in OBS is to be studied for better network throughput and protection in metropolitan networks.

REFERENCES

- [1] L.H. Bonani, F. Rudge Barbosa, E. Moschim, R. Arthur "Analysis of Electronic Buffers in Optical Packet/Burst Switching Mesh Networks", International Conference on Transport Optical Networks-ICTON-2008, June 2008 – Athens, Greece.
- [2] I. B. Martins, L. H. Bonani, F. R. Barbosa, E. Moschim, "Dynamic Traffic Analysis of Metro Access Optical Packet Switching Networks having Mesh Topologies", Proc. Int. Telecom Symp., ITS'2006, Sept. 2006, Fortaleza, Brazil.
- [3] S. Yao, B. Mukherjee, S. J. Yoo, S. Dixit, "A Unified Study of Contention Resolution Schemes in Optical packet switching Networks", IEEE J. Lightwave Tech, vol.21, no.3, p.672, March 2003.
- [4] R. Ramaswami, K.N. Sivarajan, Optical Networks: a practical perspective, Morgan Kaufmann Publishers, 2nd Edition, 2002.
- [5] I. B. Martins, L. H. Bonani, E. Moschim, F. Rudge Barbosa, "Comparison of Link failure and Protection in Ring and Mesh OPS/OBS Metropolitan Area Optical Networks", Proc 13th Symp. on Microwave and Optoelectronics- MOMAG'2008, Sept. 2008, Floripa SC, Brazil.
- [6] T. Cinkler, L. Gyarmati, "Mpp: Optimal Multi-Path Routing with Protection", proceeding Int. Conf. Communications –ICC-2008- Beijing, China.
- [7] D. A. Schupke and R. Prinz. "Capacity, Efficiency and Restorability of Path Protection and Rerouting in WDM Networks Subject to Dual Failures", Photonic Network Comm., Vol 8, n. 2, p.191, Springer, Netherlands Sept. 2004.
- [8] T. Hills. (2002) "Next-Gen SONET ", Lightreading Rep. [Online]. Available: http://www.lightreading.com/document.asp?doc_id=14781
- [9] L. Choy, "Virtual concatenation tutorial: Enhancing SONET/SDH networks for data transport," J. Opt. Networking, vol. 1, no. 1, pp. 18–29, Dec. 2001.
- [10] C. Qiao and M. Yoo, "Choices, features, and issues in optical burst switching," Opt. Network Mag., vol. 1, no. 2, pp. 36–44, 2000.
- [11] T. Battestilli and H. Perros, "An introduction to optical burst switching," IEEE Commun. Mag., vol. 41, pp. S10–S15, Aug. 2003.
- [12] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching (OBS): A new area in optical networking research," IEEE Network, to be published.
- [13] M. Duser and P. Bayvel, "Analysis of wavelength-routed optical burst-switched network performance," in Proc. Optical Communications (ECOC), vol. 1, 2001, pp. 46–47.
- [14] Sami Sheeshia, Yang Chen, Vishal Anand, "Performance Comparison of OBS and SONET in Metropolitan Ring Networks" vol. 22, no. 8, October 2004 IEEE.

Bhupesh Bhatia received the B.Tech (2000) from Maharishi Dayanand University Rohtak and M.Tech. (2004) from IASE Demed University Sardar Sahar Rajay Stan in Electronics and Communication Engineering. He is pursuing Ph.D degree from Uttarakhand Technical University Dehradun Uttarakhand His area of interest of Bhupesh Bhatia is signal&system, digital signal processing and optical fiber communication. He has a teaching experience of more than ten years. Currently he is working as Assistant Professor in Northern India Engineering College New Delhi affiliated to Guru Gobind Sing Inderprasth University New Delhi He is the author of several engineering books.

Mr.R.K.Singh received the B.Tech, M.Tech from Birla Institute of Technical Education Pilani and Ph.d. degrees Allaahabad University Alahabad from the Department of Electronic and Communication Engineering, Dr.R.K.Singh worked as a member of Academic committee of Uttarakhand Technical University Dehradun (Uttarakhand). Dr. Singh has contributed in the area of Micro electronics, Fiber optics Communication and Solid state devices. He has published several research paper in National and International Journal. He is the member of several institution and education bodies. Currently Dr.Singh working as a officer on special duty in Uttarakhand Technical University Dehradun (Uttarakhand). He is the author of several engineering books.

A Survey on Session Hijacking

P. Ramesh Babu

Dept of Computer Science & Engineering
Sri Prakash College of Engineering
Tuni-533401, INDIA

D.Lalitha Bhaskari

Dept of Computer Science & Systems Engineering
AU College of Engineering (A)
Visakhapatnam-530003, INDIA

CPVNJ Mohan Rao

Dept of Computer Science & Engineering
Avanthi Institute of Engineering & Technology
Narsipatnam-531113, INDIA

Abstract

With the emerging fields in e-commerce, financial and identity information are at a higher risk of being stolen. The purpose of this paper is to illustrate a common-cum-valiant security threat to which most systems are prone to i.e. Session Hijacking. It refers to the exploitation of a valid computer session to gain unauthorized access to information or services in a computer system. Sensitive user information is constantly transported between sessions after authentication and hackers are putting their best efforts to steal them. In this paper, we will be setting the stages for the session hijacking to occur, and then discussing the techniques and mechanics of the act of session hijacking, and finally providing general strategies for its prevention.

Key words: session hijacking, packet, application level, network level, sniffing, spoofing, server, client, TCP/IP, UDP and HTTP

1. Introduction

Session hijacking refers to the exploitation of a valid computer session to gain unauthorized access to information or services in a computer system or the session hijack is a process whereby the attacker inserts themselves into an existing communication session between two computers. Generally speaking, session hijack attacks are usually waged against a

Workstation server type of communication session; however, hijacks can be conducted between a workstation computer communicating with a network based appliance like routers, switches or firewalls. Now we will substantiate the clear view of stages and levels of session hijacking. “Indeed, in a study of 45 Web applications in production at client companies found that 31 percent of e-commerce applications were vulnerable to cookie manipulation and session hijacking” [3]. Section 2 of this paper deals with the different stages of session hijacking, section 3 deals in depth details of where session hijacking can be done followed by discussion of Avoidance of session hijacking. Section 5 concludes the paper.

2. Stages of session hijacking

Before we can discuss the details of session hijacking, we need to be familiar with the stages on which this act plays out. We have to identify the vulnerable protocols and also obtain an understanding of what sessions are and how they are used. Based on our survey, we have found that the three main protocols that manage the data flow on which session hijacking occurs are TCP, UDP, and HTTP.

2.1 TCP

TCP stands for Transmission Control Protocol. We define it as “one of the main protocols in TCP/IP networks. TCP the IP protocol deals only with packets and TCP enable two hosts to establish a connection and exchange streams of data. TCP guarantees delivery of data and also guarantees that packets will be delivered in the same order in which they were sent.”[2] The last part of TCP definition is important in our discussion of session hijacking. In order to guarantee that packets are delivered in the right order, TCP uses acknowledgement (ACK) packets and sequence numbers to create a “full duplex reliable stream connection between two end points,” [4] with the end points referring to the communicating hosts. The two figures below provide a brief description of how TCP works:

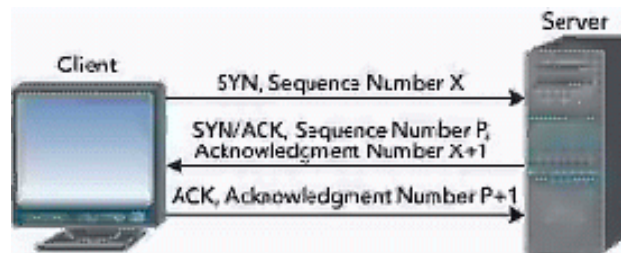


Figure 1: TCP Session establishment using Three-Way Handshake Method
(Figure and TCP summary taken from [1])

The connection between the client and the server begins with a three-way handshake (Figure 1). It proceeds as follows:

- Client sends a synchronization (SYN) packet to the server with initial sequence number X.
- Server responds by sending a SYN/ACK packet that contains the server's own sequence number p and an ACK number for the client's original SYN packet. This ACK number indicates the next

sequence number the server expects from the client.

- Client acknowledges receipt of the SYN/ACK packet by sending back to the server an ACK packet with the next sequence number it expects from the server, which in this case is P+1.

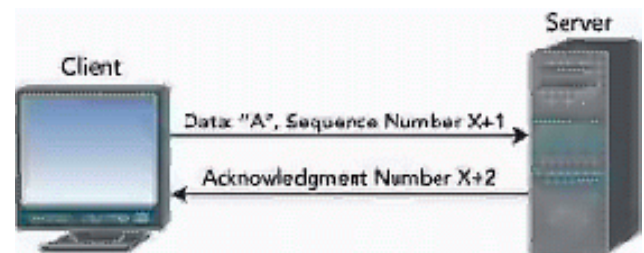


Figure 2: Sending Data over TCP
(Figure and TCP summary taken from [1])

After the handshake, it's just a matter of sending packets and incrementing the sequence number to verify that the packets are getting sent and received. In Figure 2, the client sends one byte of info (the letter "A") with the sequence number X+1 and the server acknowledges the packet by sending an ACK packet with number x+2 (x+1, plus 1 byte for the A character) as the next sequence number expected by the server. The period where all this data is being sent over TCP between client and server is called the TCP session. It is our first stage on which session hijacking will play out.

2.2 UDP

The next protocol is UDP which stands for User Datagram Protocol. It is defined as “a connectionless protocol that, like TCP, runs on top of IP networks. Unlike TCP/IP, UDP/IP provides very few error recovery services, offering instead a direct way to send and receive datagram's over an IP network.”[6] UDP doesn't use sequence numbers like TCP. It is mainly used for broadcasting messages across the network or for doing DNS queries. Online first person

shooters like Quake and Half-life make use of this protocol. Since it's connectionless and does not have any of the more complex mechanisms that TCP has, it is even more vulnerable to session hijacking. The period where the data is being sent over UDP between client and server is called the UDP session. UDP is our second stage for session hijacking.

2.3 HTTP

HTTP stands for Hyper Text Transfer Protocol. We define HTTP as the underlying protocol used by the World Wide Web. HTTP defines how messages are formatted and transmitted, and what actions Web servers and browsers should take in response to various commands. For example, when you enter a URL in your browser, this actually sends an HTTP command to the Web server directing it to fetch and transmit the requested Web page. ” [2]

It is also important to note that HTTP is a stateless protocol. Each transaction in this protocol is executed independently with no knowledge of past transactions. The result is that HTTP has no way of distinguishing one user from the next. To uniquely track a user of a web application and to persist his/her data within the HTTP session, the web application defines its own session to hold this data. HTTP is the final stage on which session hijacking occurs, but unlike TCP and UDP, the session to hijack has more to do with the web application's implementation instead of the protocol (HTTP).

3. Levels of session hijacking

Session hijacking can be done at two levels: Network Level and Application Level. Network level hijacking involves TCP and UDP sessions, whereas Application level

session hijack occurs with HTTP sessions. Attacks at each level are not unrelated, however. Most of the time, they will occur together depending on the system that is attacked. For example, a successful attack on a TCP session will no doubt allow one to obtain the necessary information to make a direct attack on the user session on the application level.

3.1 Network level hijacking

The network level refers to the interception and tampering of packets transmitted between client and server during a TCP or UDP session. Network level session hijacking is particularly attractive to hackers, because they do not have to customize their attacks on a per web application basis. It is an attack on the data flow of the protocol, which is shared by all web applications [7].

3.1.1 TCP Session hijacking

The goal of the TCP session hijacker is to create a state where the client and server are unable to exchange data, so that he can forge acceptable packets for both ends, which mimic the real packets. Thus, attacker is able to gain control of the session. At this point, the reason why the client and server will drop packets sent between them is because the server's sequence number no longer matches the client's ACK number and likewise, the client's sequence number no longer matches the server's ACK number. To hijack the session in the TCP network the hijacker should employ following techniques: they are as follows [7]

- IP Spoofing
- Blind Hijacking
- Man in the Middle attack (packet sniffing)

IP Spoofing

IP spoofing is “a technique used to gain unauthorized access to computers, whereby the intruder sends messages to a computer with an IP address indicating that the message is coming from a trusted host.”[2] Once the hijacker has successfully spoofed an IP address, he determines the next sequence number that the server expects and uses it to inject the forged packet into the TCP session before the client can respond. By doing so, he creates the “desynchronized state.” The sequence and ACK numbers are no longer synchronized between client and server, because the server registers having received a new packet that the client never sent. Sending more of these packets will create an even greater inconsistency between the two hosts.

Blind Hijacking

If source routing is disabled, the session hijacker can also employ blind hijacking where he injects his malicious data into intercepted communications in the TCP session. It is called “blind” because the hijacker can send the data or commands, but cannot see the response. The hijacker is basically guessing the responses of the client and server. An example of a malicious command a blind hijacker can inject is to set a password that can allow him access from another host.

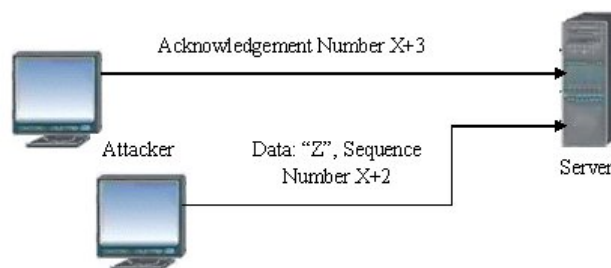


Figure3: Blind Injection

Man in the Middle attack (packet sniffing)

This technique involves using a packet sniffer that intercepts the communication between the client and server. With all the data between the hosts flowing through the hijacker's sniffer, he is free to modify the content of the packets. The trick to this technique is to get the packets to be routed through the hijacker's host. [1]

3.1.2 UDP Session hijacking

Hijacking a session over User Datagram Protocol (UDP) is exactly the same as over TCP, except that UDP attackers do not have to worry about the overhead of managing sequence number and other TCP mechanisms. Since UDP is connectionless, injecting data into session without being detected is extremely easy. If the “man in the middle” situation exists, this can be very easy for the attacker, since he can also stop the server's reply from getting to the client in the first place [6]. Figure4 shows how an attacker could do this.

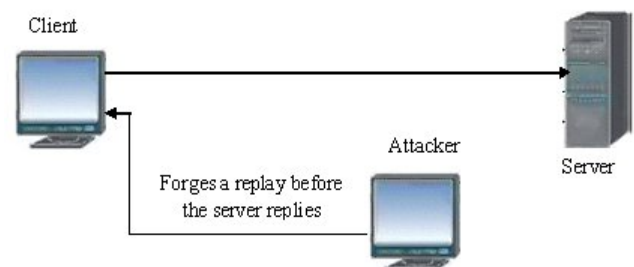


Figure4: Session Hijacking over UDP

DNS queries, online games like the Quake and Half-Life, and peer-to-peer sessions are common protocols that work over UDP; all are popular target for this kind of session hijacking.

3.2 Application level hijacking

The application level refers to obtaining session IDs to gain control of the HTTP user session as defined by the web application. In the application level, the session hijacker not only tries to hijack existing sessions, but also tries to create new sessions using stolen data. Session hijacking at the application level mainly involves obtaining a valid session ID by some means in order to gain control of an existing session or to create a new unauthorized session.

3.2.1 HTTP Session hijacking

HTTP session hijacking is all about obtaining the session ID, since web applications key off of this value to determine identity. Now we will see the techniques involved in HTTP session hijacking [7].

Obtain Session IDs

Session IDs generally can be found in three locations [5]:

- Embedded in the URL, which is received by the application through HTTP GET requests when the client clicks on links embedded with a page.
- Within the fields of a form and submitted to the application. Typically the session ID information would be embedded within the form as a hidden field and submitted with the HTTP POST command.
- Through the use of cookies.

All three of these locations are within the reach of the session hijacker. Embedded session info in the URL is accessible by looking through the browser history or proxy server or firewall logs. A hijacker can sometimes reenter in the URL from the

browser history and get access to a web application if it was poorly coded. Session info in the form submitted through the POST command is harder to access, but since it is still sent over the network, it can still be accessed if the data is intercepted. Cookies are accessible on the client's local machine and also send and receive data as the client surfs to each page. The session hijacker has a number of ways to guess the session ID or steal it from one of these locations.

Observation (Sniffing)

Using the same techniques as TCP session hijacking, the hijacker can create the "man in the middle" situation and use a packet sniffer. If the HTTP traffic is sent unencrypted, the session hijacker has traffic redirected through his host where he can examine the intercepted data and obtain the session ID. Unencrypted traffic could carry the session ID and even usernames and passwords in plain text, making it very easy for the session hijacker to obtain the information required to steal or create his own unauthorized session.

Brute Force

If the session ID appears to be predictable, the hijacker can also guess the session ID via a brute force technique, which involves trying a number of session IDs based upon the pattern. This can be easily set up as an automated attack, going through multiple possibilities until a session ID works. "In ideal circumstances, an attacker using a domestic DSL line can potentially conduct up to as many as 1000 session ID guesses per second." Therefore, if the algorithm that produces the session ID is not random enough, the session hijacker can obtain a usable session ID rather quickly using this technique.

Misdirected Trust [5]

It refers to using HTML injection and cross-site scripting to steal session information. HTML injection involves finding a way to inject malicious HTML code so that the client's browser will execute it and send session data to the hijacker. Cross-site scripting has the same goal, but more specifically exploits a web application's failure to validate user-supplied input before returning it to the client system. Cross-site" refers to the security restrictions placed on data associated with a web site (e.g. session cookies). The goal of the attack is to trick the browser into executing injected code under the same permissions as the web application domain. By doing so, he can steal session information from the client side. The success of such an attack is largely dependent on the susceptibility of the targeted web application.

4. Avoidance of Session Hijacking

To protect your network with session hijacking, a user has to implement both security measures at Application level and Network level. Network level hijacks can be prevented by ciphering the packets so that the hijacker cannot decipher the packet headers, to obtain any information which will aid in spoofing. This encryption can be provided by using protocols such as IPSEC, SSL, SSH etc. Internet security protocol (IPSEC) has the ability to encrypt the packet on some shared key between the two parties involved in communication [7]. IPSec runs in two modes: Transport and Tunnel. In Transport Mode only the data sent in the packet is encrypted while in Tunnel Mode both packet headers and data are encrypted, so it is more restrictive [4].

To prevent your Application session to be hijacked it is recommended to use

Strong Session ID's so that they cannot be hijacked or deciphered at any cost. SSL (Secure Socket layer) and SSH (Secure Shell) also provides strong encryption using SSL certificates so that session cannot be hijacked, but tools such as Cain & Bell can spoof the SSL certificates and decipher everything! Expiring sessions after a definite period of time requires re-authentication which will useless the hacker's tricks [7].

Methods to avoid session hijacking include [8]:

- An open source solution is ArpON "Arp handler inspectiON". It is a portable ARP handler which detects and blocks all Man in the Middle attacks through ARP poisoning and spoofing attacks with a static ARP inspection (SARPI) and dynamic ARP inspection (DARPI) approach on switched LANs with or without DHCP. This requires an agent on every host that is to be protected.
- Use of a long random number or string as the session key. This reduces the risk that an attacker could simply guess a valid session key through trial and error or brute force attacks.
- Regenerating the session id after a successful login. This prevents session fixation because the attacker does not know the session id of the user after he has logged in.
- Encryption of the data passed between the parties; in particular the session key. This technique is widely relied-upon by web-based banks and other e-commerce services, because it completely prevents sniffing-style attacks. However, it could still be possible to perform some other kind of session hijack.
- Some services make secondary checks against the identity of the user. For

example, a web server could check with each request made that the IP address of the user matched the one last used during that session. This does not prevent attacks by somebody who shares the same IP address, however, and could be frustrating for users whose IP address is liable to change during a browsing session.

➤ Alternatively, some services will change the value of the cookie with each and every request. This dramatically reduces the window in which an attacker can operate and makes it easy to identify whether an attack has taken place, but can cause other technical problems

➤ Users may also wish to log out of websites whenever they are finished using them

5. Conclusion

Session hijacking remains a serious threat to networks and web applications on the web. This paper provides a general overview of how the malicious exploit is done and how the information security engineer can protect networks and web applications from this threat. It is important to protect our session data at both the network and application levels. Although implementing all of the countermeasures discussed here does not completely guarantee full immunity against session hijacking, it does raise the security bar and forces the session hijacker to come up with alternate and perhaps more complex methods of attack. It is a good idea to keep testing and monitoring our networks and applications to ensure that they will not be susceptible to the hijacker's tricks.

We hope earnestly that the paper we presented will cater the needs of novice researchers and students who are interested in session hijacking.

6. References

- [1] Lam, Kevin, David LeBlanc, and Ben Smith. "Hacking: Fight Back: Theft On The Web: Prevent Session Hijacking." Microsoft TechNet Festival. Winter 2005. 1 Jan. 2005.
- [2] <<http://www.webopedia.com/>>.
- [3] Morana, Marco. "Make It and Break It: Preventing Session Hijacking and Cookie Manipulation." Secure Enterprise Summit, 23 Nov. 2004.
- [4] William Stallings, Network Security Essentials, 3rd Edition, Pearson Edition.
- [5] Ollman, Gunter, "Web Session Management: Best Practices in Managing HTTP Based Client Sessions." Technical Info: Making Sense of Security. Accessed 20 Dec. 2004.
- [6] Kevin L. Paulson, "Hack proofing your network "1st Edition, Global Knowledge Professional reference. Syngress Edition
- [7] "Session Hijacking in Windows Networks.". By Mark Lin, Date Submitted: 1/18/2005 GSEC Practical Assignment v1.4c (Option 1) of SANS Institute of Information Security.
- [8] www.wikipedia.com

Authors Profile



Ms. Dr. D. Lalitha Bhaskari is an Associate professor in the department of Computer Science and Engineering of Andhra University. She did her Phd from JNTU Hyderabad in the area of Steganography and Watermarking. Her areas of interest include Theory of computation, Data Security, Image Processing, Data communications, Pattern Recognition. Apart from her regular academic activities she holds prestigious responsibilities like Associate Member in the Institute of Engineers, Member in IEEE, Associate Member in the Pentagram Research Foundation, Hyderabad, India. She is also the recipient of “Young Engineers” Award from the prestigious Institution of Engineers (INDIA) for the year 2008 in Computer Science discipline.



Mr. P. Ramesh babu is an Assistant Professor in the Department of Computer Science & Engineering of Sri Prakash college of Engineering-Tuni. His research interests include Steganography, Digital Watermarking, Information security and Data communications. Mr. Ramesh babu did his M.Tech in Computer Science & Engineering from JNTU Kakinada. He has 5 years of good teaching experience. Contact him at: rameshbabu_kb@yahoo.co.in



Dr. C.P.V.N.J Mohan Rao is a Professor in the Department of Computer Science and Engineering and principal of Avanthi Institute of Engineering & Technology - Narsipatnam.

He did his PhD from Andhra University and his research interests include Image Processing, Networks & Data security, Data Mining and Software Engineering. He has guided more than 50 M.Tech Projects. He received many honors and he has been the member for many expert committees, member of many professional bodies and Resource person for various organizations.

Point-to-Point IM Interworking Session Between SIP and MFTS

¹Mohammed Faiz Aboalmaaly, ²Omar Amer Abouabdalla ³Hala A. Albaroodi and ⁴Ahmed M. Manasrah

National Advanced IPv6 Centre
Universiti Sains Malaysia
Penang, Malaysia

Abstract— This study introduces a new IM interworking prototype between the Session Initiation Protocol (SIP) and the Multipoint File Transfer System (MFTS). The interworking system design is presented as well. The interworking system relies on adding a new network entity to enable the interworking which has the ability to work as a SIP server to the SIP-side of the network and as a MFTS server to the MFTS-side of the network. Use Cases tool is used to describe the translation server architecture. Finally, experimental-based results show that the interworking entity is able to run a successful point-to-point interoperability IM session between SIP and MFTS that involved user registration and message translations as well.

Keywords- SIP; MFTS; Instant Messaging (IM);

I. INTRODUCTION

Over the last few years, the use of computer network systems to provide communication facilities among people has increased; hence the service provided for this area must be enhanced. Various signaling protocols have arisen and many multimedia conferencing systems have been developed that use these signaling protocols in order to provide audio, video, data and instant messaging communication among people. Transparent interoperability between dissimilar signaling protocols and Instant Messaging and Presence (IMP) applications has become desirable in order to ensure full end-to-end connectivity. In order to enable the interoperability between two or more different signaling protocols or standards, a translation mechanism must exist in between to translate the non-similar control options and media profiles. SIP [1], is a well-known signaling protocol that has been adopted in many areas and applications in the Internet as a control protocol. SIP is an application layer protocol, used for establishing, modifying and ending multimedia sessions in an IP-based network. SIP is a standard created by the Internet Engineering Task Force (IETF) for initiating an interactive user session that involves multimedia elements such as, video, voice, chat, gaming and virtual reality. It is also, a request-response protocol; like the HTTP [2], it uses messages to manage the multimedia conference over the Internet. On the other hand, The Multipoint File Transfer System or (MFTS) [3] is a file distribution system based on the well known “client-server architecture”. The MFTS server is actually a distribution engine, which handles the issues related to file sharing as well as instant messaging exchange among the various MFTS

clients. The MFTS has been adopted in the Multimedia Conferencing System (MCS) product [4] by the Document Conferencing unit (DC), which is a network component that is responsible for any user communications related to file sharing as well as instant messaging interaction.

II. SIP AND MFTS AS INSTANT MESSAGING PROTOCOLS

A. MFTS as an Instant Messaging Protocol

As everyone knows, Instant Messaging is a type of near real-time communication between two or more people based on typed text. The text is carried via devices connected over a network such as the Internet. MFTS in turn, uses control messages as a carrier to send and receive instant messages (with text) among MFTS clients. As a normal IM communication, an MFTS client sends several instant messages with a variety of lengths to one or more MFTS clients. Figure 1 depicts the standard structure of the MFTS control message

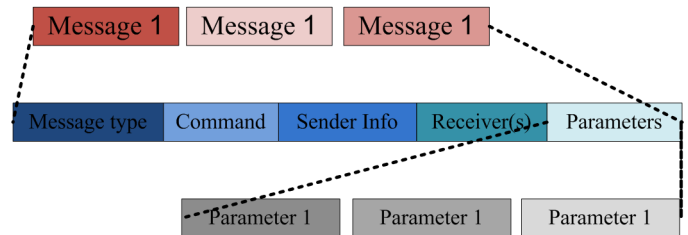


Figure 1. MFTS Message Structure

As depicted above, the MFTS message is divided into five main fields Message Type, Command, Sender Information, Receiver(s) Information, and Parameters. Message type is used to indicate the purpose of the message whether it is client to server message or it is a server to server message, while the command indicates the specific name of the message like *Private Chat* (PRCHAT), the Command is a six character length. Additionally, Sender info and receiver(s) are used to identify the IP address of both the sender and the receiver respectively. Parameters are used to identify protocol-specific issues which out of the scope of this study [5].

B. SIP as Instant Messaging Protocol

The Internet Engineering Task Force (IETF) has defined two modes of instant messaging for SIP. The first is the pager mode, which makes use of the SIP MESSAGE method, as defined in [6]. The MESSAGE method is an extension to the SIP that allows the transfer of Instant Messages. This mode establishes no sessions, but rather each MESSAGE request is sent independently and carries the content in the form of MIME (Multipurpose Internet Mail Extensions) body part of each request. Additionally, grouping these independent requests can be achieved at the SIP UA's by adding a user interface that lists these messages in ordered way or grouped in a dialog initiated by some other SIP request. By contrast, the session mode makes use of the Message Session Relay Protocol or MSRP [7], which is designed to transmit a series of related instant messages of arbitrary sizes in the context of a session.

III. INTERWORKING METHOD

As mentioned previously in [8], SIP handles two methods for instant messaging services, pager mode and session mode. In a session mode there will be a session establishment using Message Session Relay Protocol (MSRP) while in the pager mode there is no need to establish a session, because the MESSAGE method in SIP is actually a signaling message or request which is the same as INVITE, CANCEL and OPTION. On the other hand, the MFTS server is the distributing engine responsible for sending instant messages among MFTS users, which uses control messages for that purpose. From this point, we found out that it is more stable to choose the SIP pager mode for instant messaging as the other part to communicate with MFTS users. Figure 2 below shows the SIP MESSAGE request.

```
MESSAGE sip:user2@domain.com SIP/2.0
Via: SIP/2.0/TCP
user1pc.domain.com;branch=z9hG4bK776sgdkse
Max-Forwards: 70
From: sip:user1@domain.com;tag=49583
To: sip:user2@domain.com
Call-ID: asd88asd77a@1.2.3.4
CSeq: 1 MESSAGE
Content-Type: text/plain
Content-Length: 18
```

Hello World

Figure 2. SIP MESSAGE Request

Since both MFTS and SIP use the Transmission Control Protocol (TCP) for sending and receiving control messages (signaling) between their network components, the translation module should use TCP as well.

A. SIP-MFTS Interworking

In order to ensure that a message will reach its destination, SIP proxy server may forward a SIP message request to another server; in other words, a SIP message request may

traverse through several proxies before it reaches the final destination of the end user [1]. On the other hand, in MFTS, similar mechanism is used to ensure that an MFTS message will reach to the user that resided behind another MFTS [3]. The proposed interworking module will take the advantage of these features. The idea is to combine both the proxy server capabilities with MFTS server capabilities in one entity. This entity should also include a translation component that translates SIP messages to MFTS messages and vice versa. In this case, both SIP proxy server and MFTS server will communicate with this entity as a server analogous to them. Accordingly, this method will provide transparent communication to the users and to the servers as well. In addition to that, the translation process will be done within that bi-directional translation server. The Figure below illustrates the general interworking prototype between SIP and MFTS.

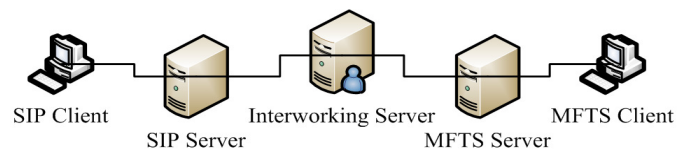


Figure 3. SIP-MFTS Interworking

B. System Model

Before starting the interworking session, the translation module must register itself with the SIP server and supports the address resolution schemes of SIP. In MFTS, there are two types of registration. The first registration is that the MFTS server should register itself to other MFTS servers, since the translation model is considered as another MFTS server from a MFTS user's side; it must register itself with MFTS server. The second type of registration is the process by which an MFTS client logs into the MFTS server, and informs it of its IP address. Registration will occur before any instant messaging sessions are attempted. The MFTS server will respond with either a confirmation or a reject message. In SIP, the REGISTER request allows a SIP registrar server to know the client's address.

C. Interworking Module Requirements

Each entity in the interworking module has been analyzed based on its normal functionalities. According to that, Figure 4 shows the internal modules by using the use case tool of the proposed translation server and the number of connections to the SIP side of the network and to the MFTS side of the network. As illustrated in figure 4, two modules are used for the registration for both SIP and MFTS, and two additional modules are used for sending and receiving the control messages, these two modules are linked together by the translation function module to translate between the two types of instant messages (MESSAGE and PRCHAT).

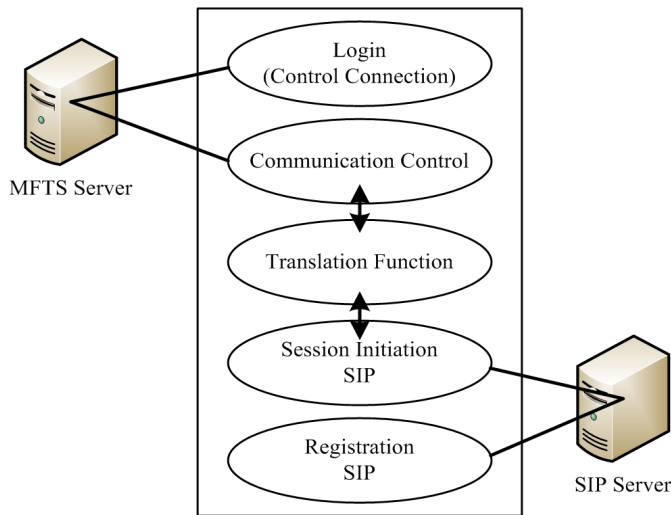


Figure 4. Use Case Diagram for the Proposed Translation Server

D. SIP and MFTS Message Translation

Both SIP and MFTS messages consist of few fields that are used to identify the sender, the receiver or receivers and some other information, in both of them this information is considered as a message header. Table I and Table II show the translation table that translates MFTS specifications to SIP specifications and from SIP specifications to MFTS specifications respectively.

TABLE I. MFTS-TO-SIP TRANSLATION TABLE

MFTS	SIP Header or Contents
Command	body of MESSAGE
Thread	Call-ID
Sender-Info	From
Receiver(s)	To

TABLE II. SIP-TO-MFTS TRANSLATION TABLE

SIP Header or contents	MFTS
Call-ID	thread
Content-Language	(no Mapping)
Cseq	(no mapping)
From	Sender-Info
Subject	(no Mapping)
To	Receiver(s)
body of MESSAGE	Command

IV. TESTING AND RESULTS

The translation server testing is based on proposing real interoperability IM scenarios. Two tests are conducted, one to check the functionality of the system as an IM interoperability module between SIP and MFTS, while the second is supplementary to the first one which is to know the time required to receive an instant message to the destination client. Both tests are applied on a one-to-one interoperability session

between SIP and MFTS. Moreover, each test is conducted five times to ensure certainty.

A. Functional Testing

SIP-MFTS Functional testing is basically done by sending several chat messages with a variety of lengths to the destination/s. It is applied on all proposed scenarios that were mentioned in subsection 5.2.1. Five different lengths of messages are sent through the network starting from "Hello world" sentence and ending with its duplications, for instance, the second sentence is "Hello world Hello world" and so on. All functional tests were done successfully.

B. Time Required

This part of testing has actually followed the same conducted steps in the functional testing. All tests at this stage are done by acquiring the required time for each chat message to reach the other domain. Furthermore, each type of test is done five times and an arithmetic mean is calculated for them. Table III reports the time required for the messages to be sent from the SIP client to the MFTS client, while Table IV shows the time required for the message to be sent from the MFTS client to the SIP client. Moreover, there was no significant difference noticed in both tests (SIP to MFTS) and (MFTS to SIP).

TABLE III. SIP TO MFTS

Message Lenght	Time (Seconds)
"Hello World" X1	0.23
"Hello World" X2	0.27
"Hello World" X4	0.34
"Hello World" X8	0.45
"Hello World" X16	0.43

TABLE IV. MFTS TO SIP

MFTS	SIP Header or Contents
"Hello World" X1	0.29
"Hello World" X2	0.28
"Hello World" X4	0.26
"Hello World" X8	0.50
"Hello World" X16	0.39

V. CONCLUSION AND FUTURE WORK

The translation server was capable of handling a one - to - one instant messaging conference between SIP and MFTS. Two types of tests were conducted; functionality test and the time required. All tests are done successfully and were within an acceptable range. Proposed future work might cover the multipoint IM sessions between SIP and MFTS (work in progress) and might also include the multiple-protocol interoperability concept that involves many IM protocols communicating together. Furthermore, since MFTS has the capability to work as a file transfer system, and since there is a study conducted to make SIP able to work as a file transfer

system based on the capability provided by MSRP, additional interworking between SIP and MFTS based on file transfer capability will increase the usefulness of this study.

REFERENCES

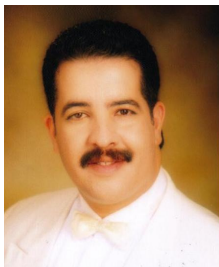
- [1] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, et al., "SIP: Session Initiation Protocol", RFC 3261, June 2002.
- [2] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, et al., "Hypertext transfer protocol-HTTP/1.1", RFC 2616, June 1999.
- [3] S. N. Saleh, "An Algorithm To Handle Reliable Multipoint File Transfer Using The Distributed Network Entities Architecture" Master Thesis, Universiti Sains Malaysia, Malaysia, 2004.
- [4] "Multimedia Conferencing System – MCS" Internet: <http://www.unimal.ac.id/mcs/MCSv6.pdf>, [17-September-2010].
- [5] B. Campbell, J. Rosenberg, H. Schulzrinne, C. Huitema, and D. Gurle, "Session Initiation Protocol (SIP) Extension for Instant Messaging", RFC 3428, December 2002.
- [6] B. Campbell, R. Mahy, C. Jennings, "The Message Session Relay Protocol (MSRP)", RFC 4975, September 2007.
- [7] S. N. Saleh, "Semi-Fluid: A Content Distribution Model For Faster Dissemination Of Data" PhD Thesis, Universiti Sains Malaysia, Malaysia, 2010.
- [8] J. C. Han, S. O. Park, S. G. Kang and H. H. Lee, "A Study on SIP-based Instant Message and Presence" in: The 9th International Conference on Advanced Communication Technology, Korea, vol 2, pp. 1298-1301, February 2007.

AUTHORS PROFILE



(MANET) and Parallel Computing.

A PhD candidate, He received his bachelor degree in software engineering from Mansour University College (IRAQ) and a master's degree in computer science from Universiti Sains Malaysia (Malaysia). His PhD. research is mainly focused on Overlay Networks. He is interested in several areas of research such as Multimedia Conferencing, Mobile Ad-hoc Network



interest include Multimedia Network, Internet Protocol version 6 (IPv6), and Network Security.

Dr. Omar Amer Abouabdalla obtained his PhD degree in Computer Sciences from University Science Malaysia (USM) in the year 2004. Presently he is working as a senior lecturer and domain head in the National Advanced IPv6 Centre – USM. He has published more than 50 research articles in Journals and Proceedings (International and National). His current areas of research



A PhD candidate joined the NAV6 in 2010. She received her Bachelor degree in computer sciences from Mansour University College (IRAQ) in 2005 and a master's degree in computer sciences from Universiti Sains Malaysia (Malaysia)

in 2009. Her PhD research is on peer-to-peer computing. She has numerous research of interest such as IPv6 multicasting and video Conferencing.



Dr. Ahmed M. Manasrah is a senior lecturer and the deputy director for research and innovation of the National Advanced IPv6 Centre of Excellence (NAV6) in Universiti Sains Malaysia. He is also the head of inetmon project "network monitoring and security monitoring platform". Dr. Ahmed obtained his Bachelor of Computer Science from Mu'tah University, al Karak, Jordan in 2002. He obtained his Master of Computer Science and doctorate from Universiti Sains Malaysia in 2005 and 2009 respectively. Dr. Ahmed is heavily involved in researches carried by NAV6 centre, such as Network monitoring and Network Security monitoring with 3 Patents filed in Malaysia.

An Extensive Survey on Gene Prediction Methodologies

Manaswini Pradhan

Lecturer, P.G. Department of Information and
Communication Technology,
Fakir Mohan University, Orissa, India

Dr. Ranjit Kumar Sahu

Assistant Surgeon, Post Doctoral Department of Plastic and
Reconstructive Surgery,
S.C.B. Medical College, Cuttack, Orissa, India

Abstract-In recent times, Bioinformatics plays an increasingly important role in the study of modern biology. Bioinformatics deals with the management and analysis of biological information stored in databases. The field of genomics is dependant on Bioinformatics which is a significant novel tool emerging in biology for finding facts about gene sequences, interaction of genomes, and unified working of genes in the formation of final syndrome or phenotype. The rising popularity of genome sequencing has resulted in the utilization of computational methods for gene finding in DNA sequences. Recently computer assisted gene prediction has gained impetus and tremendous amount of work has been carried out on this subject. An ample range of noteworthy techniques have been proposed by the researchers for the prediction of genes. An extensive review of the prevailing literature related to gene prediction is presented along with classification by utilizing an assortment of techniques. In addition, a succinct introduction about the prediction of genes is presented to get acquainted with the vital information on the subject gene prediction.

Keywords- Genomic Signal Processing (GSP), gene, exon, intron, gene prediction, DNA sequence, RNA, protein, sensitivity, specificity, mRNA.

I. INTRODUCTION

Biology and biotechnology are transforming research into an information-rich enterprise and hence they are developing technological revolution. The implementation of computer technology into the administration of biological information is Bioinformatics [3]. It is a fast growing area of computer science that deals with the collection, organization and analysis of DNA and protein sequence. Nowadays, for addressing the recognized and realistic issues which originate in the management and analysis of biological data, it incorporates the construction and development of databases, algorithms, computational and statistical methods and hypothesis [1]. It is debatable that back to Mendel's discovery of genetic inheritance in 1865, the origin of bioinformatics history can be discovered. On the other hand, bioinformatics research in a real sense began in late 1960s which is represented by Dayoff's atlas of protein sequences as well as the early modeling analysis of protein and RNA structures [3].

Due to the availability of excessive amount of genomic and proteomic data in public domain, it is becoming progressively more significant to process this information in such a way that are valuable to humankind [4]. One of the challenges in the analysis of newly sequenced genomes is the computational recognition of genes and the understanding of the genome is the fundamental step. For evaluating genomic sequences and annotate genes, it is required to discover precise and fast tools [5]. In this framework, a significant role in these fields has been played by the established and recent signal processing techniques [4]. Comparatively, Genomic signal processing (GSP) is a new field in bio-informatics that deals with the digital signal representations of genomic data and analysis of the same by means of conventional digital signal processing (DSP) techniques [6].

In the DNA (deoxyribonucleic acid) of a living organism, the genetic information is accumulated. DNA is a macro molecule in the form of a double helix. There are pairs of bases among the two strands of the backbone. There are four bases called adenine, cytosine, guanine, and thymine. They are abbreviated with the letters A, C, G, and T respectively [1]. For the chemical composition of one individual protein, Gene is a fragment of DNA consisting of the formula. Genes serve as the blueprints for proteins and a few additional products. During the production of any genetically encoded molecule, mRNA is the initial intermediate [8]. The genomic information is frequently presented by means of the sequences of nucleotide symbols in the strands of DNA molecules or by using the symbolic codons (triplets of nucleotides) or by the symbolic sequences of amino acids in the subsequent polypeptide chains [5].

Genes and the intergenic spaces are the two types of regions in a DNA sequence. Proteins are the building blocks of every organism and the information for the generation of the proteins are stored in the gene, where genes are in charge for the construction of distinct proteins. Although, every cell in an organism consists of identical DNA, only a subset is expressed in any particular family of cells and hence they have identical genes [1]. The exons and the introns are the two

regions in the genes of eukaryotes. The exons and the introns are the two regions in the genes of eukaryotes. The exons which are the protein coding region of a gene are distributed with interrupting sequences of introns. The biological significance of intron is not well known still; therefore they are termed as protein non coding regions. The borders in-between the introns and the exons are described as splice sites [9].

When a gene is expressed, it is recorded first as pre-mRNA. Then, it goes through a process called splicing where non-coding regions are eliminated. A mature mRNA which does not consist of introns, serves as a template for the synthesis of a protein in translation. In translation, each and every codon which is a collection of three adjacent base pairs in mRNA directs the addition of one amino acid to a peptide for synthesizing. Therefore, a protein is a sequence of amino acid residues subsequent to the mRNA sequence of a gene [7]. The process is shown in the fig.1,

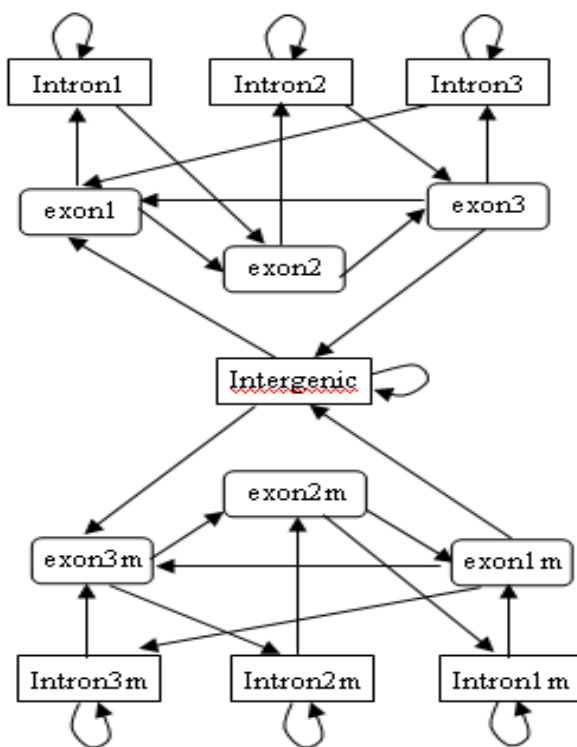


Figure 1: Transcription of RNA, splicing of intron, and translation of protein processes

One of the most important objectives of genome sequencing is to recognize all the genes. In eukaryotic genomes, the analysis of a coding region is also based on the accurate identification of the exon-intron structures. On the other hand, the task becomes very challenging due to vast length and structural complexity of sequence data. [9]. In recent years, a wide range of gene prediction techniques for

analyzing, predicting diseases and more have been reported by huge range of researchers. In this paper, we present an extensive review of significant researches on gene prediction along with its processing techniques. The prevailing literature available in gene prediction are classified and reviewed extensively and in addition we present a concise description about gene prediction. In section 2, a brief description of computational gene prediction is presented. An extensive review on the study of significant research methods in gene prediction is provided in section 3. Section 4 sums up the conclusion.

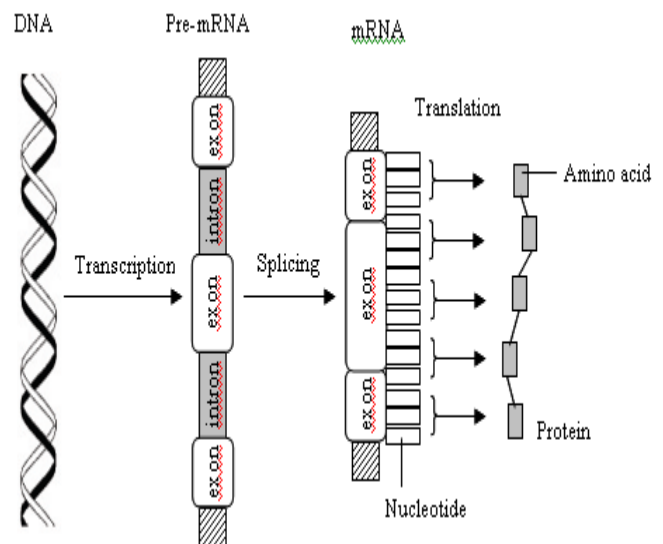


Figure 2: Gene structure's state diagram. The mirror-symmetry reveals the fact that DNA is double-stranded and genes appear on both the strands. The 3-periodicity in the state diagram correlates to the translation of nucleotide triplets into amino acids.

II. COMPUTATIONAL GENE PREDICTION

For the automatic analysis and annotation of large uncharacterized genomic sequences, computational gene prediction is becoming increasingly important [2]. Gene identification is for predicting the complete gene structure, particularly the accurate exon-intron structure of a gene in a eukaryotic genomic DNA sequence. After sequencing, finding the genes is one of the first and most significant steps in knowing the genome of a species [40]. Gene finding usually refers to the field of computational biology which is involved with algorithmically recognizing the stretches of sequence, generally genomicDNA that are biologically functional. This specially not only involves protein-coding genes but may also include additional functional elements for instance RNA genes and regulatory regions [16].

Genomic sequences which are constructed now are with length in the order of many millions of base pairs. These sequences contain a group of genes that are separated from each other by long stretches of intergenic regions [10]. With the intention of providing tentative annotation on the location,

structure and the functional class of protein-coding genes, the difficulty in gene identification is the problem of interpreting nucleotide sequences by computer [13]. The improvement of techniques for identifying the genes in DNA sequences and for genome analysis, evaluating their functions is significant [12].

Almost 20 years ago, gene identification efforts have been started and it constructed a huge number of practically effectual systems [11]. In particular, this not only includes protein-coding genes but also additional functional elements for instance RNA genes and regulatory regions. Calculation of protein-coding genes includes identification of correct splice and translation of signals in DNA sequences [14]. On the other hand, due to the exon-intron structure of eukaryotic genes, prediction is problematical. Introns are the non-coding regions that are spliced out at acceptor and donor splice sites [17].

Gene prediction is used for involving prediction of genes proteins [15]. The gene prediction accurateness is calculated using the standard measures, sensitivity and specificity. For a feature for instance coding base, exon and gene, the sensitivity is the number of properly predicted features that are separated by the number of annotated features. The specificity is defined as the number of appropriately predicted features alienated by the number of predicted features. A predicted exon is measured correct if both the splice sites are at annotated position of an exon. A predicted gene is measured correct if all the exons are properly predicted and there should be no additional exons in the annotation. Predicted partial genes were estimated as predicted genes [10]. The formulas for sensitivity and specificity are shown below.

Sensitivity: The fraction of identified genes (or bases or exons) which are correctly predicted.

$$S_n = \frac{TP}{\text{all true in reality}} = \frac{TP}{TP + FN}$$

where TP - True Positive, FN - False Negative

Specificity: The fraction of predicted genes (or bases or exons) which corresponds to true genes

$$S_p = \frac{TP}{\text{all true in prediction}} = \frac{TP}{TP + FP}$$

III. EXTENSIVE REVIEW OF SIGNIFICANT RESEARCHES ON GENE PREDICTION

A wide range of research methodologies employed for the analysis and the prediction is presented in this section. The reviewed gene prediction based on some mechanisms are classified and detailed in the following subsections.

A. Support Vector Machine

Jiang Qian *et al.* [70] presented an approach which depends upon the SVMs for predicting the targets of a transcription factor by recognizing subtle relationships between their expression profiles. Particularly, they used SVMs for predicting the regulatory targets for 36 transcription factors in the *Saccharomyces cerevisiae* genome which depends on the microarray expression data from lots of different physiological conditions. In order to incorporate an important number of both positive and negative examples, they trained and tested their SVM on a data set that are constructed by discussing the data imbalance issues directly. This was non-trivial where nearly all the known experimental information specified is only for positives. On the whole, they discovered that 63% of their TF-target relationships were approved by means of cross-validation. By analyzing the performance with the results from two recent genome-wide ChIP-chip experiments, they further estimated the performance of their regulatory network identifications. On the whole, the agreement between their results and those experiments which can be comparable to the agreement (albeit low) between the two experiments have been discovered by them. With a specified transcription factor having targets comparatively broaden evenly over the genome, they identified that this network has a delocalized structure regarding the chromosomal positioning.

MicroRNAs (miRNAs) which play an important role as post transcriptional regulators are small non-coding RNAs. For the 5' components, the purpose of animal miRNAs normally depends upon complementarities. Even though lot of suggested numerous computational miRNA target-gene prediction techniques, they still have drawbacks in revealing actual target genes. MiTarget which is a SVM classifier for miRNA target gene prediction have been introduced by Kim *et al.* [38]. As a similarity measure for SVM features, it used a radial basis function kernel and is then classified by structural, thermodynamic, and position-based features. For the first time, it presented the features and it reproduced the mechanism of miRNA binding. With the help of biologically relevant data set that is achieved from the literature, the SVM classifier has created high performance comparing with earlier tools. Using Gene Ontology (GO) analysis, they calculated important tasks for human miR-1, miR-124a, and miR-373 and from a feature selection experiment, explained the importance of pairing at positions 4, 5, and 6 in the 5' region of a miRNA. They have also presented a web interface for the program.

A Bayesian framework depends upon the functional taxonomy constraints for merging the multiple classifiers have been introduced by Zafer Barutcuoglu *et al.* [67]. A hierarchy of SVM classifiers has been trained on multiple data types. For attaining the most probable consistent set of predictions, they have merged predictions in the suggested Bayesian framework. Experiments proved that the suggested Bayesian

framework has enhanced predictions for 93 nodes over a 105-node sub-hierarchy of the GO. Accurate positioning of SVM margin outputs to probabilities has also been provided by their technique as an added advantage. They have completed function predictions for multiple proteins using this method and they approved the predictions for proteins that are involved in mitosis by experiments.

Alashwal *et al.* [19] represented Bayesian kernel for the Support Vector Machine (SVM) in order to predict protein-protein interactions. By integrating the probability characteristic of the existing experimental protein-protein interactions data, the classifier performances which were compiled from different sources could be enhanced. Besides to that, in order to organize more research on the highly estimated interactions, the biologists are boosted with the probabilistic outputs which are achieved from the Bayesian kernel. The results have implied that by using the Bayesian kernel compared to the standard SVM kernels, the accuracy of the classifier has been improved. Those results have suggested that by using Bayesian kernel, the protein-protein interaction could be computed with better accuracy as compared to the standard SVM kernels.

B. Gene ontology

A method for approximating the protein function from the Gene Ontology classification scheme for a subset of classes have been introduced by Jensen *et al.* [73] This subset which incorporated numerous pharmaceutically appealing categories such as transcription factors, receptors, ion channels, stress and immune response proteins, hormones and growth factors can be calculated. Even though the method depended on protein sequences as the sole input, it did not depend on sequence similarity. Instead it relied on the sequence derived protein features for instance predicted post translational modifications (PTMs), protein sorting signals and physical/chemical properties predicted from the amino acid composition. This granted prediction of the function for orphan proteins in which not a single homologs can be achieved. They recommended two receptors in the human genome using this method and in addition they confirmed chromosomal clustering of related proteins.

Hongwei Wu *et al.* [42] introduced a computational method for predicting the functional modules which are encoded in microbial genomes. They have also acquired a formal measure for measuring the degree of consistency among the predicted and the known modules and carried out statistical analysis of consistency measures. From three different perspectives such as phylo genetic profile analysis, gene neighborhood analysis and Gene Ontology assignments, they firstly estimated the functional relationship between two genes. Later, they combined the three different sources of information in the framework of Bayesian inference and by using the combined information; they computed the strength of gene functional relationship. Lastly, they applied a threshold-based method for

predicting the functional modules. They predicted 185 functional modules by executing this method to *Escherichia coli* K12. In *E.coli*, their estimation was extremely reliable with the previously known functional modules. The application results have confirmed that the suggested approach shows high potential for determining the functional modules which are encoded in a microbial genome.

Ontology-based pattern identification (OPI) is a data mining algorithm that methodically recognizes expression patterns that best symbolizes on hand information of gene function. Rather than depending on a widespread threshold of expression resemblance to describe functionally connected sets of genes, OPI obtained the optimal analysis background that produce gene expression patterns and gene listings that best predict gene function utilizing the criterion of GBA. Yingyao Zhou *et al.* [58] have utilized OPI to a publicly obtainable gene expression data collection on the different stages of life of the malarial parasite *Plasmodium falciparum* and methodically annotated genes for 320 practical types on the basis of existing Gene Ontology annotations. An ontology-based hierarchical tree of the 320 types gave a systems-wide biological perspective of this significant malarial parasite.

Remarkable advancement in sequencing technology and sophisticated experimental assays that interrogate the cell, along with the public availability of the resulting data, indicate the era of systems biology. There is an elemental obstacle for development in system biology as the biological functions of more than 40% of the genes in sequenced genomes remain unidentified. The development of techniques that can automatically make use of these datasets to make quantified and robust predictions of gene function that are experimentally verified require comprehensive and wide variety of available data. The VIRtual Gene Ontology (VIRGO) introduced by Massjouni *et al.* [35]. They have described that a functional linkage network (FLN) is build upon from gene expression and molecular interaction data and these genes are labeled in the FLN with their functional annotations in their Gene Ontology and these labels are systematically propagated across the FLN in order to specifically predict the functions of unlabelled genes. The helpful supplementary data for evaluating the quality of the predictions and prearranging them for further analysis was provided by the VIRGO. The survival of gene expression data and functional annotations in other organisms makes the expanding of VIRGO effortless in them. An informative 'propagation diagram' was provided for every prognosis by the VIRGO to sketch the course of data in the FLN that led to the prediction.

Important approach into the cellular function and machinery of a proteome has been provided using a map of protein-protein interactions. With a relative specificity semantic relation, the similarity between two Gene Ontology (GO) terms is measured. Here, a method for restructuring a yeast protein-protein interaction map that exclusively depends upon the GO observations has been presented by Wu *et al.*

[37]. Using high-quality interaction datasets, this technique has been confirmed for its efficiency. A positive dataset and a negative dataset for protein-protein interactions, based on a Z-score analysis were acquired. Additionally, a gold standard positive (GSP) dataset which has the highest level of confidence covered 78% of the high-quality interaction dataset and a gold standard negative (GSN) dataset which has the lowest level of confidence were acquired. Additionally, using the positives and the negatives as well as GSPs and GSNs, they determined four high-throughput experimental interaction datasets. Their supposed network which consists of 40 753 interactions among 2259 proteins has been regenerated from GSPs and configured 16 connected components. Apart from homodimers onto the predicted network, they defined every MIPS complex. Consequently, 35% of complexes were recognized to be interconnected. They also recognized few non-member proteins for seven complexes which may be functionally associated to the concerned complexes.

The functions of each protein are performed inside some specialized locations in a cell. For recognizing the protein function and approving its purification, this subcellular location is important. For predicting the location which depends upon the sequence analysis and database information from the homologs, there are numerous computational techniques. Few latest methods utilize text obtained from biological abstracts. The main goal of Alona Fyshe *et al.* [72] is to enhance the prediction accuracy of such text-based techniques. For improving text-based prediction, they recognized three techniques such as (1) a rule for ambiguous abstract removal, (2) a mechanism for using synonyms from the Gene Ontology (GO) and (3) a mechanism for using the GO hierarchy to generalize terms. They proved that these three methods can enhance the accuracy of protein sub-cellular location predictors considerably which utilized the texts that are removed from PubMed abstracts whose references were preserved in Swiss-Prot.

C. Homology

Chang *et al.* [21] introduced a scheme for improving the accuracy of gene prediction that has merged the ab-initio method based on homology. Taking the advantage of the known information, the latter recognizes each gene for previously recognized genes whereas, the former rely on predefined gene features. In spite of the crucial negative aspect of the homology-based method, the proposed scheme has also adopted parallel processing for assuring the optimal system performance i.e. the bottleneck happened predictably due to the large amount of unprocessed ordered information.

Automatic gene prediction is one of the predominant confrontations in computational sequence analysis. Conventional methods to gene detection depend on statistical models derived from already known genes. Contrary to this, a set of comparative methods depend on likening genomic sequences from evolutionary associated organisms to one another. These methods were founded on the hypothesis of

phylogenetic foot printing: they capitalize on the feature that functionally significant areas in genomic sequences are generally more conserved than non-functional areas. Taher *et al.* [53] have constructed a web-based computer program for gene prediction on the basis of homology at BiBiServ (Bielefeld Bioinformatics Server). The input data given to the tool is a duo of evolutionary associated genomic sequences e.g., from human and mouse. The server runs CHAOS and DIALIGN to produce an arrangement of the input sequences and later searched for the conserved splicing indicators and start/stop codons in the neighborhood areas of local sequence conservation. Genes were predicted on the basis of local homology data and splice indicators. The server submitted the predicted genes along with a graphical representation of the fundamental arrangement.

Perfect accuracy is yet to be attained in computational gene prediction techniques, even for comparatively simple prokaryotic genomes. Problems in gene prediction revolve around the fact that several protein families continue to be uncharacterized. Consequently, it appears that only about half of an organism's genes can be assuredly ascertained on the basis of similarity with other known genes. Hossain Sarker *et al.* [46] have attempted to discern the intricacies of certain gene prediction algorithms in Genomics. Furthermore, they have attempted to discover the advantages and disadvantages of those algorithms. Ultimately, they have proposed a new method for Splice Alignment Algorithm that takes into account the merits and demerits of it. They anticipated that the proposed algorithm will subdue the intricacies of the existing algorithm and ensure more precision.

D. Hidden Markov Model (HMM)

Pavlovic *et al.* [20] have presented a well organized framework in order to learn the combination of gene prediction systems. Their approach can model the statistical dependencies of the experts which is the main advantage. The application of a family of combiners has been represented by them in the increasing order of statistical complexity starting from a simple Naïve Bayes to Input HMMs. A system has been introduced by them for combining the predictions of individual experts in a frame-consistent manner. This system depends on the stochastic frame consistency filter which is implemented as a Bayesian network in the post-combination stage. Intrinsically, the application of expert combiners has been enabled by the system for general gene prediction. The experiments predicted that while generating a frame-consistent decision, the system has drastically enhanced concerning the best single expert. They have also experimented that the suggested approach was in principle applicable to other predictive tasks for instance promoter or transcription elements recognition.

The computational method which was introduced for the problem of finding the genes in eukaryotic DNA

sequences is not yet solved acceptably. Gene finding programs have accomplished comparatively high accuracy on short genomic sequences but do not execute well if there is a presence of long sequences of indefinite number of genes. Here, programs which exist tend to calculate many false exons. For the ab initio prediction of protein coding genes in eukaryotic genomes a program named AUGUSTUS has been introduced by Stanke *et al.* [27]. Based on the Hidden Markov Model, the program was constructed and it incorporated a number of well-known methods and submodels. It has employed a way of modeling intron lengths. They have used a donor splice site model which directly upstream for a short region of the model that takes the reading frames into account. Later, they have applied a method which has allowed better GC-content dependent parameter estimation. Comparing AUGUSTUS which predicted that human and drosophila genes on longer sequences are far more accurate than the ab initio gene prediction programs while being more specific at the same time.

The presence of processed pseudogenes: nonfunctional, intronless copies of real genes found elsewhere in the genome damaged the correct gene prediction. The processed pseudogenes are usually mistaken for real genes or exons by gene prediction programs which lead to biologically irrelevant gene predictions. Despite the fact that the methods exists for identifying the processed pseudogenes in genomes, there has not been made any attempt for incorporating pseudogene removal with gene prediction or even for providing a freestanding tool which identifies such incorrect gene predictions. PPFINDER (for Processed Pseudogene finder), a program that has been incorporated with numerous methods of processed pseudogene for finding the mammalian gene annotations have been introduced by Van Baren *et al.* [39]. For removing the pseudogenes from N-SCAN gene predictions, they used PPFINDER and demonstrated that when gene prediction and pseudogene masking were interleaved, the gene prediction has been enhanced considerably. Additionally, they utilized PPFINDER with gene predictions as a parent database by eradicating the need for libraries of known genes. This has permitted them to manage the gene prediction/PPFINDER procedure on the newly sequenced genomes for which few genes were known.

DeCaprio *et al.* [33] demonstrated the first proportional gene predictor, Conrad which depends upon semi-Markov conditional random fields (SMCRFs). In contradictory to the best standalone gene predictors that depends upon generalized hidden Markov models (GHMMs) and accustomed by maximum probability Conrad was favourably trained for maximizing annotation accuracy. Added to this, Conrad encoded all sources of information as features and treated all features equally in the training and inference algorithms, unlike the best annotation pipelines, entrusted on heuristic and ad hoc decision rules to combine standalone gene predictors with additional information such as ESTs and protein homology. Conrad excels the best

standalone gene predictors in cross-validation and whole chromosome testing on two fungi with hugely different gene structures. SMCRF's discriminative training methods and their capability to effortlessly integrate different types of data by encoding them as feature functions gives better performance. Effectiveness of Twinscan was intimately synchronized to the duplication of prognosis of a two-species phylo-GHMM by integrating Conrad on *Cryptococcus neoformans*. Allowing discriminative training and accumulating feature functions increase the efficiency in order to acquire a level of accuracy unparalleled for their organism. While correlating Conrad versus Fgenesh on *Aspergillus nidulans* same results are obtained. Their exceedingly modular nature makes SMCRF a hopeful agenda for gene prediction by simplifying the process of designing and testing potential indicators of gene structure. SMCRFs improved the condition of the art in gene prediction in fungi by the accomplishment of Conrad's and it provides a healthy platform.

The majority of computational tools which exists depend on sequence homology and/or structural similarity for discovering microRNA (miRNA) genes. Of late, with regards to sequence, structure and comparative genomics information, the supervised algorithms were applied for addressing this problem. Almost in these studies, experimental evidence rarely supported miRNA gene predictions. In addition to, prediction accuracy remains uncertain. In order to predict the miRNA precursors, a computational tool (SSCprofiler) which utilized a probabilistic method based on Profile Hidden Markov Models was introduced by Oulas *et al.* [28]. SSCprofiler has attained a performance accuracy of 88.95% sensitivity and 84.16% specificity on a large set of human miRNA genes using the concurrent addition of biological features such as sequence, structure and conservation. The novel miRNA gene candidates situated within cancer-associated genomic regions, the trained classifier has been used for recognizing and ranking the resulting predictions using the expression information from a full genome tiling array. Lastly, using northern blot analysis, four of the top scoring predictions were confirmed by experimentation. Their work combined both analytical and experimental techniques for demonstrating that SSCprofiler which can be used to recognize novel miRNA gene candidates in the human genome was a highly accurate tool.

E. Different Software programs for gene prediction

A computational technique to create gene models by utilizing evidence produced from a varied set of sources, inclusive of those representatives of a genome annotation pipeline has been detailed by Allen *et al.* [51]. The program, known as Combiner, took into account genomic sequence as input and the positions of gene predictions from ab initio gene locators, protein sequence arrangements, expressed sequence tag and cDNA arrangements, splice site predictions, and other proofs. Three diverse algorithms for merging proof in the Combiner were realized and checked on 1783 verified genes

in *Arabidopsis thaliana*. Their results have proved that merging gene prediction proofs always excelled even the most excellent individual gene locator and, in certain cases, can create dramatic enhancements in sensitivity and specificity.

Issac *et al.* [52] have detailed that EGPred is an internet-based server that united ab initio techniques and similarity searches to predict genes, specifically exon areas, with high precision. The EGPred program consists of the following steps: (1) a preliminary BLASTX search of genomic sequence across the RefSeq database has been utilized to find protein hits with an $E - value < 1$; (2) a second BLASTX search of genomic sequence across the hits from the preceding run with relaxed parameters ($E - values < 10$) assists to get back all possible coding exon regions; (3) a BLASTN search of genomic sequence across the intron database was then utilized to identify possible intron regions; (4) the possible intron and exon regions were likened to filter/remove incorrect exons; (5) the NNSPLICE program was then utilized to relocate splicing signal site locations in the outstanding possible coding exons; and (6) ultimately ab initio predictions were united with exons obtained from the fifth step on the basis of the relative strength of start/stop and splice signal regions as got from ab initio and similarity search. The combination method augmented the exon level achievement of five diverse ab initio programs by 4%–10% when assessed on the HMR195 data set. Analogous enhancement was noticed when ab initio programs were assessed on the Burset/Guigo data set. Ultimately, EGPred has been verified on a ~95-Mbp section of human chromosome 13. The EGPred program is computationally strenuous because of multiple BLAST runs in each analysis.

Zhou *et al.* [43] introduced a gene prediction program named GeneKey. GeneKey can attain the high prediction accuracy for genes with moderate and high C+G contents when the widely used dataset which are collected by Kulp and Reese are trained [45]. On the other hand, the prediction accuracy was lesser for CG-poor genes. They constructed a LCG316 dataset which composes of gene sequences with low C+G contents to solve this problem. When the CG-poor genes are trained with LCG316 dataset, the prediction accuracy of GeneKey has been enhanced significantly. Additionally, the statistical analysis confirmed that some structure features for instance splicing signals and codon usage of CG-poor genes somewhat differ from that of CG-rich ones. GeneKey is enabled by combining the two datasets to achieve high and balanced prediction accuracy for both CG-rich and CG-poor genes. The results of their work have suggested that or enhancing the performance of different prediction tasks, careful construction of training dataset was very significant.

Mario Stanke *et al.* [48] have presented an internet server for the computer program AUGUSTUS, which is utilized to predict genes in eukaryotic genomic sequences. AUGUSTUS is founded on a comprehensive hidden Markov model representation of the probabilistic model of a sequence and its gene structure. The web server has permitted the user

to enforce constraints on the calculated gene structure. A constraint can indicate the location of a splice site, a translation commencement site or a stop codon. Moreover, it is practicable to indicate the location of acknowledged exons and gaps that were acknowledged to be exonic or intronic sequence. The number of constraints was optional and constraints can be joined in order to locate larger elements of the predicted gene structure. The outcome would be the most expected gene structure that conformed with all specified user constraints, if such a gene structure was present. The specification of constraints is helpful when portion of the gene structure is identified, e.g. by expressed sequence tag or protein sequence arrangements, or if the user wishes to alter the default prediction.

Overall of 143 prokaryotic genomes were achieved with an efficient version of the prokaryotic genefinder EasyGene. By Comparing the GenBank and RefSeq annotations with the EasyGene predictions, they unveiled that in some genomes up to 60% of the genes might be represented with an incorrect initial codon particularly in the GC-rich genomes. The fractional differentiation between annotated and predicted affirmed that numerous short genes are annotated in numerous organisms. Additionally, there is a chance that genes might be left behind during the annotation of some of the genomes. Out of 143, 41 genomes to be over-annotated by .5% which means that too many ORFs were represented as genes have been calculated by Pernille Nielsen *et al.* [68]. They also confirmed that 12 of 143 genomes were under-annotated. These results depended upon the difference between the number of annotated genes that are not found by EasyGene and the number of predicted genes that are not annotated in GenBank. They defended that the average performance of their consistent and entirely automated method was some extent improved than the annotation.

Starcevic *et al.* [31] has accomplished the program package 'ClustScan' (Cluster Scanner) for rapid, semi-automatic, annotation of DNA sequences encoding modular biosynthetic enzymes that consists of polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS) and hybrid (PKS / NRPS) enzymes. In addition of displaying the predicted chemical structures of products the program also allows the export of the structures in a standard format for analyses with other programs. Topical advancement in realizing the enzyme function has been integrated to make knowledge-based prognosis concerning the stereochemistry of products. The easy assimilation of additional knowledge regarding domain specificities and function has been allowed by the program structure. Using a graphical interface the results of analyses were offered to the user and it also allowed trouble-free editing of the predictions to acquire user experience. Annotation of biochemical pathways in microbial, invertebrate animal and metagenomic datasets demonstrate the adaptability of their program package. The annotation of all PKS and NRPS clusters in a complete Actinobacteria genome in 2–3 man hours was allowed by the speed and convenience

of the package. The easy amalgamation with other programs and promoting additional analyses of results was allowed by the open architecture of ClustScan that were valuable for a wide range of researchers in the chemical and biological sciences.

Kai Wang *et al.* [56] have built up a committed, publicly obtainable, splice site prediction program known as NetAspGene, for the genus *Aspergillus*. Gene sequences from *Aspergillus fumigatus*, the most general mould pathogen, were utilized to construct and experiment their model. Compared to several animals and plants, *Aspergillus* possesses finer introns; consequently they have utilized a bigger window dimension on single local networks for instruction, to encompass both donor and acceptor site data. They have utilized NetAspGene to remaining *Aspergilli*, including *Aspergillus nidulans*, *Aspergillus oryzae*, and *Aspergillus niger*. Assessment with unrelated data sets has exposed that NetAspGene executed considerably better splice site prediction compared to other existing tools. NetAspGene is very useful for the analysis in *Aspergillus* splice sites and specifically in alternative splicing.

The ease of use of a huge part of the maize B73 genome sequence and originating sequencing technologies recommend economical and simple ways to sequence areas of interest from many other maize genotypes. Gene content prediction is one of the steps required to convert these sequences into valuable data. Gene predictor specifically trained for maize sequences is so far not available in public. The EuGene software merged numerous sources of data into a condensed gene model prediction and this EuGene is preferred for training by Pierre Montalent *et al.* [66]. The results were compacted together into a library file and e-mailed to the user. The library includes the parameters and options utilized for predicting; the submitted sequence, the masked sequence (if relevant), the annotation file (gff, gff3 and fasta format) and a HTML file which permitted the results to be displayed by a web browser.

F. Other Training methodologies

Huiqing Liu *et al.* [69] introduced a computational method for patient outcome prediction. In the training phase of this method, they utilized two types of extreme patient samples: (1) short-term survivors who got an inconvenient result in a small period and (2) long-term survivors who were preserving a positive outcome after a long follow-up time. A clear platform has been generated for by these tremendous training samples for recognizing suitable genes whose expression was intimately related to the outcome. In order to construct a prediction model, the chosen extreme samples and the significant genes were then incorporated with the help of a support vector machine. Using that prediction model, each validation sample is allocated a risk score that falls into one of the special pre-defined risk groups. This method has been adapted by them to several public datasets. In several cases as seen in their Kaplan–Meier curves, patients in high and low

risk groups which are graded by the suggested method have evidently apparent outcome status. They have also proved that for improving the prediction accuracy, the suggestion of choosing only extreme patient samples for training is effective when different gene selection methods are utilized.

According to the parent of origin, Imprinted genes are epigenetically modified genes whose expression can be determined. They are concerned in embryonic development and imprinting dysregulation is linked to diabetes, obesity, cancer and behavioral disorders such as autism and bipolar disease. A statistical model which depends on DNA sequence characteristics have been trained by Herein, Luedi *et al.* [45]. It not only identified potentially imprinted genes but also predicted the parental allele from which they were expressed. Out of 23,788 interpreted autosomal mouse genes, their model has recognized 600 (2.5%) to be imprinted substantially, 64% of which has been estimated for revealing maternal expression. The predictions which are allowed for the recognition of putative candidate genes for complicated situations where parent-of-origin effects are involved, includes Alzheimer disease, autism, bipolar disorder, diabetes, male sexual orientation, obesity, and schizophrenia. From the experiments, it has been proved that the number, type and relative orientation of repeated elements flanking a gene are on the whole significant for predicting whether a gene was imprinted.

G. Other Machine Learning Techniques

Seneff *et al.* [24] described an approach incorporating constraints from orthologous human genes in order to predict the exon-intron structures of mouse genes using the techniques which are utilized in speech and natural language processing applications in the past. A context-free grammar is used in their approach for parsing a training corpus of annotated human genes. For capturing the common features of a mammalian gene, a statistical training process has generated a weighted Recursive Transition Network (RTN). This RTN has been extended into a finite state transducer (FST) and composed with an FST to capture the specific features of the human ortholog. The recommended model includes a trigram language model on the amino acid sequence as well as exon length constraints. For aligning the top N candidates in the search space, a final stage has used CLUSTALW which is a free software package. They have attained 96% sensitivity and 97% specificity at the exon level on the mouse genes for a set of 98 orthologous human-mouse pairs where only given knowledge are accumulated from the annotated human genome.

An approach to the problem of splice site prediction, by applying stochastic grammar inference was presented by Kashiwabara *et al.* [49]. Four grammar inference algorithms to infer 1465 grammars were used, and a 10-fold cross-validation to choose the best grammar for every algorithm was also used. The matching grammars were entrenched into a classifier and

the splice site prediction was made to run and the results were compared with those of NNSPLICE, the predictor used by Genie gene finder. Possible paths to improve this performance were indicated by using Sakakibara's windowing technique to discover probability thresholds that will lower false positive prediction.

Hoff *et al.* [26] introduced a gene prediction algorithm for metagenomic fragments based on a two-stage machine learning approach. In the first step, for extracting the features from DNA sequences, they have used linear discriminants for monocodon usage, dicodon usage and translation initiation sites. In the second step, for computing the probability in such a way that the open reading frame encodes a protein, an artificial neural network combined these features with open reading frame length and fragment GC-content. For categorizing and attaining the gene candidates, this probability was used. On artificially fragmented genomic DNA, their method produced fast single fragment predictions with good quality sensitivity and specificity by means of extensive training. In addition to that, this technique can accurately calculate translation initiation sites and differentiate the complete genes from incomplete genes with high consistency. For predicting the genes in metagenomic DNA fragments, extensive machine learning methods were compatible. Especially, the association of linear discriminants and neural networks was very promising and are supposed to be considered for incorporating into metagenomic analysis pipelines.

Single nucleotide polymorphisms (SNPs) give much assurance as a source for disease-gene association. However, the cost of genotyping the tremendous number of SNPs restricted the research. Therefore, for identifying a small subset of informative SNPs, the supposed tag SNPs is of much importance. This subset comprises of chosen SNPs of the genotypes, and represents the rest of the SNPs accurately. Additionally, in order to estimate prediction accuracy of a set of tag SNPs, an efficient estimation method is required. A genetic algorithm (GA) to tag SNP problems, and the K-nearest neighbor (K-NN) which act as a prediction method of tag SNP selection have been applied by Chuang *et al.* [23]. The experimental data which is used consists of genotype data rather than haplotype data and was taken from the HapMap project. The recommended method consistently identifies the tag SNPs with significantly better prediction accuracy than those methods from the literature. Concurrently, the number of tag SNPs which was recognized is smaller than the number of tag SNPs identified in the other methods. When the matching accuracy was reached, it is observed that the run time of the recommended method was much shorter than the run time of the SVM/STSA method.

H. Digital Signal Processing

The protein-coding areas of DNA sequences have been noticed to display the period-three behaviour, which can

be capitalized on to predict the position of coding areas inside genes. Earlier, discrete Fourier transform (DFT) and digital filter-based techniques have been utilized for the detection of coding areas. But, these techniques do not considerably subdue the noncoding areas in the DNA spectrum at $2\pi/3$. As a result, a non-coding area may unintentionally be recognized as a coding area. Trevor W. Fox *et al.* [55] have set up a method (a quadratic window operation subsequent to a single digital filter operation) that has restrained almost each of the non-coding areas. They have offered a technique that needs only one digital filter operation subsequent to a quadratic windowing operation. The quadratic window yielded a signal that has approximately zero energy in the non-coding areas. The proposed technique can be thus enhances the probability of properly recognizing coding areas over earlier digital filtering methods. Nevertheless, the precision of the proposed technique was affected when handling coding areas that do not display strong period-three behavior.

The basic problem to interpret genes is to predict the coding regions in large DNA sequences. For solving that problem, Digital Signal Processing techniques have been used successfully. Furthermore, the existing tools are not able to calculate all the coding regions which are present in a DNA sequence. A predictor introduced by Fuentes *et al.* [5] based on the linear combination of two other methods proved good quality efficacy separately. And also for reducing the computational load, a fast algorithm was developed [25] earlier. Some thoughts have been reviewed concerning the combination of the predictor with other methods. Compared to the previous methods, the efficiency of the suggested predictor was estimated by using ROC curves which showed improved performance in the detection of coding regions. The comparison in terms of computation time in between the Spectral Rotation Measure using the direct method and the proposed predictor using the fast algorithm confirmed that the computational load did not increase considerably even when the two predictors are combined.

Several digital signal processing, methods have been utilized to mechanically differentiate protein coding areas (exons) from non-coding areas (introns) in DNA sequences. Mabrouk *et al.* [57] have differentiated these sequences in relation to their nonlinear dynamical characteristics, for example moment invariants, correlation dimension, and biggest Lyapunov exponent estimates. They have utilized their model to several real sequences encrypted into a time series utilizing EIIP sequence indicators. To differentiate between coding and non coding DNA areas, the phase space trajectory was initially rebuilt for coding and non-coding areas. Nonlinear dynamical characteristics were obtained from those areas and utilized to examine a difference between them. Their results have signified that the nonlinear dynamical features have produced considerable dissimilarity between coding (CR) and non-coding areas (NCR) in DNA sequences. Ultimately, the classifier was experimented on real genes where coding and non-coding areas are widely known.

Genomic sequence, structure and function analysis of various organisms has been a testing problem in bioinformatics. In this context protein coding region (exon) identification in the DNA sequence has been accomplishing immense attention over a few decades. By exploiting the period-3 property present in it these coding regions can be recognized. The discrete Fourier transform has been normally used as a spectral estimation technique to extract the period-3 patterns available in DNA sequence. The conventional DFT approach loses its efficiency in case of small DNA sequences for which the autoregressive (AR) modeling is used as an optional tool. An optional but promising adaptive AR method for the similar function has been proposed by Sahu *et al.* [22]. Simulation study that has been done on various DNA sequences subsequently exposed that a substantial savings in computation time is accomplished by our techniques without debasing the performance. The potentiality of the planned techniques has been authenticated by means of receiver operating characteristic curve (ROC) analysis.

1. Neural Network

Alistair M. Chalket *et al.* [79] have presented a neural network based computational model that uses a broad range of input parameters for AO (Antisense Oligonucleotides) prediction. From AO scanning experiments in the literature sequence and efficacy data were gathered and a database of 490 AO molecules was generated. A neural network model was trained utilizing a set of parameters derived on the basis of AO sequence properties. On the whole a correlation coefficient of 0.30 ($p = 10^{-8}$) was obtained by the best model consisting of 10 networks. Effective AOs (>50% inhibition of gene expression) can be predicted by their model with a success rate of 92%. On an average 12 effective AOs were predicted by their model out of 1000 pairs utilizing these thresholds, thus making it an inflexible but practical method for AO prediction.

Takatsugu Kan *et al.* [75] have aimed to detect the candidate genes involved in lymph node metastasis of esophageal cancers, and investigate the possibility of using these gene subsets in artificial neural networks (ANNs) analysis for estimating and predicting occurrence of lymph node metastasis. With 60 clones their ANN model was capable of most accurately predicting lymph node metastasis. For lymph node metastasis, the highest predictive accuracy of ANN in recently added cases that were not utilized by SAM for gene selection is 10 of 13 (77%) and in all cases it is 24 of 28 (86%) (sensitivity: 15/17, 88%; specificity: 9/11, 82%). The predictive accuracy of LMS was 9 of 13 (69%) in recently added cases and 24 of 28 (86%) in all cases (sensitivity: 17/17, 100%; specificity: 7/11, 67%). It is hard to extract relevant information by clustering analysis for the prediction of lymph node metastasis.

In bioinformatics identification of short DNA sequence motifs which act as binding targets for transcription factors is an important and challenging task. Though unsupervised learning techniques are often applied from the literature of statistical theory, for the discovery of motif in large genomic datasets an effective solution is not yet found. For motif-finding problem, Shaun Mahony *et al.* [76] have offered three self-organizing neural networks. The core system SOMBRERO is a SOM-based motif-finder. The generalized models for structurally related motifs are automatically constructed and the SOMBRERO is initialized with relevant biological knowledge by the SOM-based method to which the motif-finder is integrated. Also the relationships between various motifs were displayed by a self-organizing tree method and it was proved that an effective structural classification is possible by such a method for novel motifs. By utilizing various datasets, they have evaluated the performance of the three self organizing neural networks.

Neural networks are long time popular approaches for intelligent machines development and knowledge discovery. Nevertheless, problems such as fixed architecture and excessive training time still exist in neural networks. This problem can be solved by utilizing the neuro-genetic approach. Neuro-genetic approach is based on a theory of neuroscience which states that the genome structure of the human brain considerably affects the evolution of its structure. Therefore the structure and performance of a neural network is decided by a gene created. Assisted by the new theory of neuroscience, Zainal A. Hasibuan *et al.* [77] have proposed a biologically more reasonable neural network model to overcome the existing neural network problems by utilizing a simple Gene Regulatory Network (GRN) in a neuro-genetic approach. A Gene Regulatory Training Engine (GRTE) has been proposed by them to control, evaluate, mutate and train genes. After that, based on the genes from GRTE a distributed and Adaptive Nested Neural Network (ANNN) was constructed to handle uncorrelated data. Evaluation and validation was accomplished by conducting experiments using Proben1's Gene Benchmark Datasets. The experimental results confirmed the objective of their proposed work.

Liu Qicai *et al.* [78] have employed Artificial Neural Networks (ANN) for analyzing the fundamental data obtained from 78 pancreatitis patients and 60 normal controls consisting of three structural of HBsAg, ligand of HBsAg and clinical immunological characterizations, laboratory data and genotypes of cationic trypsinogen gene PRSS1. They have verified the outcome of ANN prediction using T-cell culture with HBV and flow cytometry. The characteristics of T-cells competent of existing together with the secreted HBsAg in patients with pancreatitis were analyzed utilizing T-cell receptor from A121T, C139S, silent mutation and normal PRSS1 gene. To verify that HBsAg-specific T-cells receptor is affected by the PRSS1 gene a comparison was made on the rate of multiplication and CD4/CD8 of T-cell after culture with HBV at 0H, 12H, 24H, 36H, 48H and 72H time point.

The protein's structural predicted by the ANN was capable of identifying specific turbulence and differences of anti-HBs lever of the pancreatitis patients. One suspected HBsAg-specific T-cell receptor is the three-dimensional of the protein present with the PRSS1 gene that corresponds to HBsAg. T-cell culture has produced different results for different genotypes of PRSS1. Silent mutation and normal controls groups are considerably lower than that of PRSS1 mutation (A121T and C139S) in T-cell proliferation as well as CD4/CD8.

J. On other techniques

Rice xa5 gene produces recessive, race-specific impediment to bacterial blight disease attributable to the pathogen *Xanthomonas oryzae* pv. *Oryzae* and has immense importance for research and propagation. In an attempt to clone xa5, an F2 population of 4892 individuals was produced by Yiming *et al.* [44], from the xa5 close to isogenic lines, IR24 and IRBB5. A fine mapping process was performed and strongly linked RFLP markers were utilized to filter a BAC library of IRBB56, a defiant rice line having the xa5 gene. A 213 kb contig encompassing the xa5 locus was created. Consistent with the sequences from the International Rice Genome Sequencing Project (IRGSP), the Chinese Super hybrid Rice Genome Project (SRGP) and certain sub-clones of the contig, twelve SSLP and CAPS markers were created for precise mapping. The xa5 gene was mapped to a 0.3 cM gap between markers K5 and T4, which covered a span of roughly 24 kb, co-segregating with marker T2. Sequence assay of the 24 kb area showed that an ABC transporter and a basal transcription factor (TFIIa) were prospective candidates for the xa5 defiant gene product. The molecular system by which the xa5 gene affords recessive, race-specific resistance to bacterial blight is explained by the functional experiments of the 24 kb DNA and the candidate genes.

Gautam Aggarwal *et al.* [62] analyzed the interpretation of three complete genomes by means of the ab initio methods of gene identification GeneScan and GLIMMER. The interpretation made by means of GeneMark is endowed in GenBank which is the standard against which these are compared. In addition to the number of genes anticipated by both proposed methods, they also found a number of genes anticipated by GeneMark, but they are not identified by both of the non-consensus methods they used. The three organisms considered were the entire prokaryotic species having reasonably compact genomes. The source for a proficient non-consensus method for gene prediction is created by the Fourier measure and the measure was utilized by the GeneScan algorithm. Three complete prokaryotic genomes were used to benchmark the program and the GLIMMER. For entire genome analysis, many attempts are made to study the limitations of the recommended techniques. As long as gene-identification is involved, GeneScan and GLIMMER are of analogous accurateness with sensitivities and specificities generally higher than 0.9. GeneScan and GLIMMER

techniques provide similar results in a significant number of cases but usually the number of false predictions (both positive and negative) was higher for GeneScan than GLIMMER. It is recommended that there are some unrevealed additional genes in these three genomes and also some of the reputed identifications made previously might need re-evaluation.

Freudenberg *et al.* [64] introduced a technique for predicting disease related human genes from the phenotypic emergence of a query disease. Corresponding to their phenotypic similarity diseases of known genetic origin are to be clustered. Every cluster access includes a disease and its basic disease gene. In these clusters, recognizing the disease genes, which were phenotypically related to the query disease, were secured by the functional similarity of the potential disease genes from the human genome. Leave-one-out cross-validation of 878 diseases from the OMIM database, by means of 10672 candidate genes from the human genome is used to implement the computation of the recommended approach. Based on the functional specification, the true solution is enclosed within the top scoring 3% of predictions roughly in one-third of the cases and the true solution is also enclosed within the top scoring 15% of the predictions in two-third of the cases. The results of prognosis are used to recognize target genes, when probing for a mutation in monogenic diseases or for selection of loci in genotyping experiments in genetically complex diseases.

Thomas Schiex *et al.* [60] have detailed the FrameD, a program that predicts the coding areas in prokaryotic and matured eukaryotic sequences. In the beginning intended at gene prediction in bacterial GC affluent genomes, the gene model utilized in FrameD also permits predicting genes in the existence of frame shifts and partly undetermined sequences which makes it also remarkably appropriate for gene prediction and frame shift correction in uncompleted sequences for example EST and EST cluster sequences. Similar to current eukaryotic gene prediction programs, FrameD also has the capability to consider protein resemblance information in its prediction as well as in its graphical output. Its functioning were assessed on diverse bacterial genomes

Rice xa5 gene produces recessive, race-specific impediment to bacterial blight disease attributable to the pathogen *Xanthomonas oryzae* pv. *Oryzae* and has immense importance for research and propagation. In an attempt to clone xa5, an F2 population of 4892 individuals was produced by Yiming *et al.* [61], from the xa5 close to isogenic lines, IR24 and IRBB5. A fine mapping process was performed and strongly linked RFLP markers were utilized to filter a BAC library of IRBB56, a defiant rice line having the xa5 gene. A 213 kb contig encompassing the xa5 locus was created. Consistent with the sequences from the International Rice Genome Sequencing Project (IRGSP), the Chinese Super hybrid Rice Genome Project (SRGP) and certain sub-clones of

the contig, twelve SSLP and CAPS markers were created for precise mapping. The xa5 gene was mapped to a 0.3 cM gap between markers K5 and T4, which covered a span of roughly 24 kb, co-segregating with marker T2. Sequence assay of the 24 kb area showed that an ABC transporter and a basal transcription factor (TFIIa) were prospective candidates for the xa5 defiant gene product. The molecular system by which the xa5 gene affords recessive, race-specific resistance to bacterial blight is explained by the functional experiments of the 24 kb DNA and the candidate genes.

Bayesian variable choosing for prediction utilizing a multinomial probit regression model with data amplification to change the multinomial problem into a series of smoothing problems has been dealt with by Zhou *et al.* [50]. There are more than one regression equations and they have sought to choose the same fittest genes for all regression equations to compose a target predictor set or, in the perspective of a genetic network, the dependency set for the target. The probit regressor is estimated as a linear association of the genes and a Gibbs sampler has been engaged to determine the fittest genes. Numerical methods to hurry up the calculation were detailed. Subsequent to determining the fittest genes, they have predicted the destination gene on the basis of the fittest genes, with the coefficient of determination being utilized to evaluate predictor precision. Utilizing malignant melanoma microarray data, they have likened two predictor models, the evaluated probit regressors themselves and the optimal entire logic predictor on the basis of the chosen fittest genes, and they have likened these to optimal prediction not including feature selection. Some rapid implementation issues for this Bayesian gene selection technique have been detailed, specifically, calculating estimation errors repeatedly utilizing QR decomposition. Experimental results utilizing malignant melanoma data has proved that the Bayesian gene selection gives predictor sets with coefficients of determination that are competent with those got via a complete search across all practicable predictor sets.

A reaction pattern library which consists of bond-formation patterns of GT reactions have been introduced by Shin Kawano *et al.* [71] and the co-occurrence frequencies of all reaction patterns in the glycan database is researched. Using this library and a co-occurrence score, the prediction of glycan structures was pursued. In the prediction method, a penalty score was also executed. Later, using the individual reaction pattern profiles in the KEGG GLYCAN database as virtual expression profiles, they examined the presentation of prediction by means of the leave-one-out cross validation method. 81% was the accuracy of prediction. Lastly, the real expression data have applied to the prediction method. Glycan structures consists of sialic acid and sialyl Lewis X epitope which were predicted by use of the expression profiles from the human carcinoma cell, concurred well with experimental outcomes.

A comparative-based method to the gene prediction issue has been offered by Adi *et al.* [47]. It was founded on a syntenic arrangement of more than two genomic sequences. In other words, on an arrangement that took into account the truth that these sequences contain several conserved regions, the exons, interconnected by unrelated ones, the introns and intergenic regions. To the creation of this alignment, the predominant idea was to excessively penalize the mismatches and intervals within the coding regions and inappreciably penalize its occurrences within the non-coding regions of the sequences. This altered type of the Smith-Waterman algorithm has been utilized as the foundation of the center star approximation algorithm. With syntenic arrangement they indicated an arrangement that was made considering the feature that the involved sequences contain conserved regions interconnected by unconserved ones. This method was realized in a computer program and verified the validity of the method on a standard containing triples of human, mouse and rat genomic sequences on a standard containing three triples of single gene sequences. The results got were very encouraging, in spite of certain errors detected for example prediction of false positives and leaving out of small exons.

Linkage analysis is a successful process for combining the diseases with particular genomic regions. These regions are usually big, incorporating hundreds of genes that make the experimental methods engaged to recognize the disease gene arduous and cost. In order to prioritize candidates for more experimental study, George *et al.* [40] have introduced two techniques: Common Pathway Scanning (CPS) and Common Module Profiling (CMP). CPS depends upon the supposition that general phenotypes are connected with dysfunction in proteins which contribute in the same complex or pathway. CPS implemented the network data that are derived from the protein-protein interaction (PPI) and pathway databases for recognizing associations between genes. CMP has recognized similar candidates using a domain-dependent sequence similarity approach depending upon the assumption that interruption of genes of identical function may direct to the similar phenotype. Both algorithms make use of two forms of input data namely known disease genes and multiple disease loci. When known disease genes is used as input, the combination of both techniques have a sensitivity of 0.52 and a specificity of 0.97 and it decreased the candidate list by 13-fold. Using multiple loci, their suggested techniques have recognized the disease genes for every benchmark diseases successfully with a sensitivity of 0.84 and a specificity of 0.63.

For deciphering the digital information that is stored in the human genome, the most important goal is to identify and characterize the complete ensemble of genes. Many algorithms have been described for computational gene predictions which are eventually resulted from two fundamental concepts likely modeling gene structure and recognizing sequence similarity. Successful hybrid methods combining these two concepts have also been developed. A

third orthogonal approach for gene prediction which depends on the detection of the genomic signatures of transcription have been introduced by Glusman *et al.* [41] and are accumulated over evolutionary time. Depending upon this third concept, they have considered four algorithms: Greens and CHOWDER which calculates the mutational strand biases that are caused by transcription-coupled DNA repair and ROAST and PASTA which are based on strand-specific selection against polyadenylation signals. Aggregating these algorithms into an incorporated method called FEAST; they anticipated the location and orientation of thousands of putative transcription units not overlapping known genes. Several previously predicted transcriptional units did not arrived for coding the proteins. The recent algorithms are mainly suitable for the detection of genes with lengthy introns and that lack sequence conservation. Therefore, they have accomplished the existing gene prediction methods and helped for identifying the functional transcripts within various apparent “genomic deserts”.

Differing from most organisms, the α -proteobacterium *Acidithiobacillus ferrooxidans* withstand an abundant supply of soluble iron and they live in dreadfully acidic conditions (pH 2). It is also odd that it oxidizes iron as an energy source. Therefore, it faces the demanding twin problems of managing intracellular iron homeostasis when accumulated with enormously elevated environmental masses of iron and modifying the utilization of iron both as an energy source and as a metabolic micronutrient. Recognizing Fur regulatory sites in the genome of *A. ferrooxidans* and to gain insight into the organization of its Fur regulon are undergone by a combination of bioinformatic and experimental approach. Wide range of cellular functions comprising metal trafficking (e.g. *feoPABC*, *tdr*, *tonBexbBD*, *copB*, *cdf*), utilization (e.g. *fdx*, *nif*), transcriptional regulation (e.g. *phoB*, *irr*, *iscR*) and redox balance (*grx*, *trx*, *gst*) that are connected by fur regulatory targets is identified. FURTA, EMSA and in vitro transcription analyses affirmed the anticipated Fur regulatory sites. The first model for a Fur-binding site consensus sequence in an acidophilic iron-oxidizing microorganism was given by Quatrini *et al.* [34] and he laid the foundation for forthcoming studies aimed at expanding their understanding of the regulatory networks that control iron uptake, homeostasis and oxidation in extreme acidophiles.

A generic DNA microarray design which suits to any species would significantly benefit comparative genomics. The viability of such a design by ranking the great feature densities and comparatively balanced nature of genomic tiling microarrays was proposed by Royce *et al.* [36]. In particular, first of all, they separated every *Homo sapiens* Refseq-derived gene's spliced nucleotide sequence into all possible contiguous 25 nt subsequences. Then for each and every 25 nt subsequences, they have investigated a modern human transcript mapping experiment's probe design for the 25 nt probe sequence which have the smallest number of mismatches with the subsequence, however that did not match

the subsequence accurately. For predicting the gene expression levels in each and every experiment's thirty-three hybridizations, signal intensities which measured with each and every gene's nearest-neighbor features were equated consequently. In terms of both sensitivity and specificity, they inspected the fidelity of the suggested approach in order to detect actively transcribed genes for transcriptional consistency among exons of the identical gene and for reproducibility between tiling array designs. Overall, their results presented proof-of-principle for searching nucleic acid targets with off-target, nearest-neighbor features.

For analyzing the functional gene links, the phylogenetic approaches have been compared by Daniel Barker *et al.* [74]. From species' genomes, the independent instances of the correlated gain and loss of pairs of genes have been encountered by using these approaches. They interpreted the effect from the significant results of correlations on two phylogenetic approaches such as Dollo parsimony and maximum likelihood (ML). They investigated further the consequence which limits the ML model by setting up the rate of gene gain at a low value rather than approximating from the data. With a case study of 21 eukaryotic genomes and test data that are acquired from known yeast protein complexes, they recognized the correlated evolution among a test set of pairs of yeast (*Saccharomyces cerevisiae*) genes. During the detection of known functional links, ML acquired the best results considerably, only when the rate of the genes which were gained was controlled to low. Later, the model had smaller number of parameters but it was more practical to restrict genes from being gained more than once.

The complex and restrained problem in eukaryotes is accurate gene prediction. A constructive feature of predictable distributions of spliceosomal intron lengths were presented by William Roy *et al.* [32]. Intron lengths were not anticipated to respect coding frame as the introns were detached from transcripts prior to translation. Consequently, the number of genomic introns which are a manifold of three bases ('3n introns') must be analogous to the number that were a multiple of three plus one bases (or plus two bases). The significance of skews in intron length distributions suggests the methodical errors in intron prediction. Occasionally a genome-wide surfeit of 3n introns suggest that several internal exonic sequences are incorrectly called introns, whereas a discrepancy of 3n introns suggest that numerous 3n introns that lack stop codons are mistaken for exonic sequence. The skew in intron length distributions was shown as a general problem from the analysis of genomic interpretation for 29 diverse eukaryotic species. It is considered that the specific problem with gene prediction was specified by several examples of skews in genome-wide intron length distribution. It is recommended that a rapid and easy method for disclosing a selection of probable methodical biases in gene prediction or even problems with genome assemblies is the assessment of length distributions of predicted introns and it is also well

thought-out the ways in which these insights could be integrated into genome annotation protocols.

Poonam Singhal *et al.* [59] have introduced an ab initio model for gene prediction in prokaryotic genomes on the basis of physicochemical features of codons computed from molecular dynamics (MD) simulations. The model necessitates a statement of three computed quantities for each codon, the double-helical trinucleotide base pairing energy, the base pair stacking energy, and a codon propensity index for protein-nucleic acid interactions. Fixing these three parameters, for every codon, facilitates the computation of the magnitude and direction of a cumulative three-dimensional vector for any length DNA sequence in all the six genomic reading frames. Analysis of 372 genomes containing 350,000 genes has proved that the orientations of the gene and non-gene vectors were considerably apart and a clear dissimilarity was made possible between genic and non-genic sequences at a level comparable to or better than presently existing knowledge-based models trained based on empirical data, providing a strong evidence for the likelihood of a unique and valuable physicochemical classification of DNA sequences from codons to genomes.

Manpreet Singh *et al.* [54] have detailed that the drug invention process has been commenced with protein identification since proteins were accountable for several functions needed for continuance of life. Protein recognition further requires the identification of protein function. The proposed technique has composed a categorizer for human protein function prediction. The model utilized a decision tree for categorization process. The protein function has been predicted based on compatible sequence derived characteristics of each protein function. Their method has incorporated the improvement of a tool which identifies the sequence derived features by resolving various parameters. The remaining sequence derived characteristics are identified utilizing different web based tools.

The efficiency of their suggested approach in type 1 diabetes (T1D) was examined by Gao *et al.* [63]. While organizing the T1D base, 266 recognized disease genes and 983 positional candidate genes were obtained from the 18 authorized linkage loci of T1D. Even though their high network degrees ($p < 1e-5$) are regulated it is found that the PPI network of recognized T1D genes have discrete topological features from others with extensively higher number of interactions among themselves. They characterized those positional candidates which are the first degree PPI neighbors of the 266 recognized disease genes to be the new candidate disease genes. This resulted in further study of a list of 68 genes. Cross validation by means of the identified disease genes as benchmark revealed that the enrichment is ~ 17.1 folded over arbitrary selection, and ~ 4 folded better than using the linkage information alone. After eliminating the co-citation with the recognized disease genes, the citations of the fresh candidates in T1D-related publications were found to

be considerably ($p < 1e-7$) greater than random and they were considerably over-represented ($p < 1e-10$) in the top 30 GO terms experienced by known disease genes. Besides, the sequence analysis exposed that they enclosed appreciably ($p < 0.0004$) greater protein domains that they were known to be applicable to T1D. Indirect validation of the recently predicted candidates has been produced by these results.

A de novo prediction algorithm for ncRNA genes with factors resulting from sequences and structures of recognized ncRNA genes in association to allure was illustrated by Thao T. Tran *et al.* [65]. Bestowing these factors, genome-wide prediction of ncRNAs was performed in *Escherichia coli* and *Sulfolobus solfataricus* by administering a trained neural network-based classifier. The moderate prediction sensitivity and specificity of 68% and 70% respectively in their method is used to identify windows with potential for ncRNA genes in *E.coli*. They anticipated 601 candidate ncRNAs and reacquired 41% of recognized ncRNAs in *E.coli* by relating windows of different sizes and with positional filtering strategies. They analytically explored six candidates by means of Northern blot analysis and established the expression of three candidates namely one represented by a potential new ncRNA, one associated with stable mRNA decay intermediates and one the case of either a potential riboswitch or transcription attenuator caught up in the regulation of cell division. Normally, devoid of the requirement of homology or structural conservation, their approach facilitated the recognition of both cis- and transacting ncRNAs in partially or completely sequenced microbial genomes.

A comparative-based method to the gene prediction issue has been offered by Adi *et al.* [30]. It was founded on a syntenic arrangement of more than two genomic sequences. In other words, on an arrangement that took into account the truth that these sequences contain several conserved regions, the exons, interconnected by unrelated ones, the introns and intergenic regions. To the creation of this alignment, the predominant idea was to excessively penalize the mismatches and intervals within the coding regions and inappreciably penalize its occurrences within the non-coding regions of the sequences. This altered type of the Smith-Waterman algorithm has been utilized as the foundation of the center star approximation algorithm. With syntenic arrangement they indicated an arrangement that was made considering the feature that the involved sequences contain conserved regions interconnected by unconserved ones. This method was realized in a computer program and verified the validity of the method on a standard containing triples of human, mouse and rat genomic sequences on a standard containing three triples of single gene sequences. The results got were very encouraging, in spite of certain errors detected for example prediction of false positives and leaving out of small exons.

MicroRNAs (miRNAs) that control gene expression by inducing RNA cleavage or translational inhibition are small

noncoding RNAs. Most human miRNAs are intragenic and they are interpreted as a part of their hosting transcription units. The gene expression profiles of miRNA host genes and their targets which are correlated inversely have been assumed by Gennarino *et al.* [29]. They have developed a procedure named HOCTAR (host gene oppositely correlated targets), which ranks the predicted miRNA target genes depending upon their anti-correlated expression behavior comparing to their respective miRNA host genes. For monitoring the expression of both miRNAs (through their host genes) and candidate targets, HOCTAR was the means for miRNA target prediction systematically that put into use the same set of microarray experiments. By applying the procedure to 178 human intragenic miRNAs, they found that it has performed better than existing prediction softwares. The high-scoring HOCTAR predicted targets which were reliable with earlier published data, were enhanced in Gene Ontology categories, as in the case of miR-106b and miR-93. Using over expression and loss-of-function assays, they have also demonstrated that HOCTAR was proficient in calculating the novel miRNA targets. They have identified its efficiency by using microarray and qRT-PCR procedures, 34 and 28 novel targets for miR-26b and miR-98, respectively. On the whole, they have alleged that the use of HOCTAR reduced the number of candidate miRNA targets drastically which are meant for testing are compared with the procedures which exclusively depends on target sequence recognition.

IV. DIRECTIONS FOR THE FUTURE RESEARCH

In this review paper, various techniques utilized for the gene prediction has been analyzed thoroughly. Also, the performance claimed by the technique has also been analyzed. From the analysis, it can be understood that the prediction of genes using the hybrid techniques shown the better accuracy. Due to this reason, the hybridization of more techniques will attain the acute accuracy in prediction of genes. This paper will be a healthier foundation for the budding researchers in the gene prediction to be acquainted with the techniques available in it. In future lot of innovative brainwave will be rise using our review work

V. CONCLUSION

Gene prediction is a rising research area that has received growing attention in the research community over the past decade. In this paper, we have presented a comprehensive survey of the significant researches and techniques existing for gene prediction. An introduction to gene prediction has also been presented and the existing works are classified according to the techniques implemented. This survey will be useful for the budding researchers to know about the numerous techniques available for gene prediction analysis.

REFERENCES

- [1] Cassian Strassle and Markus Boos, "Prediction of Genes in Eukaryotic DNA", Technical Report, 2006
- [2] Wang, Chen and Li, "A brief review of computational gene prediction methods", Genomics Proteomics, Vol.2, No.4, pp.216-221, 2004
- [3] Rabindra Ku.Jena, Musbah M.Aqel, Pankaj Srivastava, and Prabhat K.Mahanti, "Soft Computing Methodologies in Bioinformatics", European Journal of Scientific Research, Vol.26, No.2, pp.189-203, 2009
- [4] Vaidyanathan and Byung-Jun Yoon, "The role of signal processing concepts in genomics and proteomics", Journal of the Franklin Institute, Vol.341, No.2, pp.111-135, March 2004
- [5] Anibal Rodriguez Fuentes, Juan V. Lorenzo Ginori and Ricardo Grau Abalo, "A New Predictor of Coding Regions in Genomic Sequences using a Combination of Different Approaches", International Journal of Biological and Life Sciences, Vol. 3, No.2, pp.106-110, 2007
- [6] Achuth Sankar S. Nair and MahaLakshmi, "Visualization of Genomic Data Using Inter-Nucleotide Distance Signals", In Proceedings of IEEE Genomic Signal Processing, Romania, 2005
- [7] Rong she, Jeffrey Shih-Chieh Chu, Ke Wang and Nansheng Chen, "Fast and Accurate Gene Prediction by Decision Tree Classification", In Proceedings of the SIAM International Conference on Data Mining,, Columbus, Ohio, USA, April 2010
- [8] Anandhavalli Gauthaman, "Analysis of DNA Microarray Data using Association Rules: A Selective Study", World Academy of Science, Engineering and Technology, Vol.42, pp.12-16, 2008
- [9] Akma Baten, Bch Chang, Sk Halgamuge and Jason Li, "Splice site identification using probabilistic parameters and SVM classification", BMC Bioinformatics, Vol.7, No.5, pp.1-15, December 2006
- [10] Te-Ming Chen, Chung-Chin Lu and Wen-Hsiung Li, "Prediction of Splice Sites with Dependency Graphs and Their Expanded Bayesian Networks", Bioinformatics, Vol.21, No.4, pp.471-482, 2005
- [11] Nakata, Kanchesia and Delisi, "Prediction of splice junctions in mRNA sequences", Nucleic Acids Research, Vol.14, pp.5327-5340, 1985
- [12] Shigehiko Kanaya, Yoshihiro Kudo, Yasukazu Nakamura and Toshimichi Ikemura, "Detection of genes in Escherichia coli sequences determined by genome projects and prediction of protein production levels, based on multivariate diversity in codon usage", Cabios, Vol.12, No.3, pp.213-225, 1996
- [13] Fickett, "The gene identification problem: an overview for developers", Computers and Chemistry, Vol.20, No.1, pp.103-118, March 1996
- [14] Axel E. Bernal, "Discriminative Models for Comparative Gene Prediction", Technical Report, June, 2008
- [15] Ying Xu and peter Gogarten, "Computational methods for understanding bacterial and archaeal genomes", Imperial College Press, Vol.7, 2008
- [16] Skarlas Lambrosa, Ioannidis Panos and Likiothanassis Spiridona, "Coding Potential Prediction in Wolbachia Using Artificial Neural Networks", Silico Biology, Vol.7, pp.105-113, 2007
- [17] Igor B.Rogozin, Luciano Milanese and Nikolay A. Kolchanov, "Gene structure prediction using information on homologous protein sequence", Cabios, Vol.12, No.3, pp.161-170, 1996
- [18] Joel H. Graber, "computational approaches to gene finding", Report, The Jackson Laboratory, 2009
- [19] Hany Alashwal, Safaai Deris and Razib M. Othman, "A Bayesian Kernel for the Prediction of Protein-Protein Interactions", International Journal of Computational Intelligence, Vol. 5, No.2, pp.119-124, 2009
- [20] Vladimir Pavlovic, Ashutosh Garg and Simon Kasif, "A Bayesian framework for combining gene predictions", Bioinformatics, Vol.18, No.1, pp.19-27, 2002
- [21] Jong-won Chang, Chungoo Park, Dong Soo Jung, Mi-hwa Kim, Jae-woo Kim, Seung-sik Yoo and Hong Gil Nam, "Space-Genes : Microbial Gene Prediction System Based on Linux Clustering", Genome Informatics, Vol.14, pp.571-572, 2003.
- [22] Sitanshu Sekhar Sahu and Ganapati Panda, "A DSP Approach for Protein Coding Region Identification in DNA Sequence", International Journal of Signal and Image Processing, Vol.1, No.2, pp.75-79, 2010
- [23] Li-Yeh Chuang, Yu-Jen Hou and Cheng-Hong Yang, "A Novel Prediction Method for Tag SNP Selection using Genetic Algorithm based on KNN", World Academy of Science, Engineering and Technology, Vol.53, No.213, pp.1325-1330, 2009
- [24] Stephanie Seneff, Chao Wang and Christopher B.Burge, "Gene structure prediction using an orthologous gene of known exon-intron structure", Applied Bioinformatics, Vol.3, No.2-3, pp.81-90, 2004

- [25] Fuentes, Ginori and Abalo, "Detection of Coding Regions in Large DNA Sequences Using the Short Time Fourier Transform with Reduced Computational Load," LNCS, vol.4225, pp. 902-909, 2006.
- [26] Katharina J Hoff, Maïke Tech, Thomas Lingner, Rolf Daniel, Burkhard Morgenstern and Peter Meinicke, "Gene prediction in metagenomic fragments: A large scale machine learning approach", BMC Bioinformatics, Vol. 9, No.217, pp.1-14, April 2008.
- [27] Mario Stanke and Stephan Waack, "Gene prediction with a hidden Markov model and a new intron submodel", Bioinformatics Vol. 19, No. 2, pp.215-225, 2003
- [28] Anastasis Oulas, Alexandra Boutla, Katerina Gkirtzou, Martin Reczko, Kriton Kalantidis and Panayiota Poirazi, "Prediction of novel microRNA genes in cancer-associated genomic regions-a combined computational and experimental approach", Nucleic Acids Research, Vol.37, No.10, pp.3276-3287, 2009
- [29] Vincenzo Alessandro Gennarino, Marco Sardiello, Raffaella Avellino, Nicola Meola, Vincenza Maselli, Santosh Anand, Luisa Cutillo, Andrea Ballabio and Sandro Banfi, "MicroRNA target prediction by expression analysis of host genes", Genome Research, Vol.19, No.3, pp.481-490, March 2009
- [30] Chengzhi Liang, Long Mao, Doreen Ware and Lincoln Stein, "Evidence-based gene predictions in plant genomes", Genome Research, Vol.19, No.10, pp.1912-1923, 2009
- [31] Antonio Starcevic, Jurica Zucko, Jurica Simunkovic, Paul F. Long, John Cullum and Daslav Hranueli, "ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures", Nucleic Acids Research, Vol.36, No.21, pp.6882-6892, October 2008
- [32] Scott William Roy and David Penny, "Intron length distributions and gene prediction", Nucleic Acids Research, Vol.35, No.14, pp.4737-4742, 2007
- [33] David DeCaprio, Jade P. Vinson, Matthew D. Pearson, Philip Montgomery, Matthew Doherty and James E. Galagan, "Conrad: Gene prediction using conditional random fields", Genome Research, Vol.17, No.9, pp.1389-1398, August 2007
- [34] Raquel Quatrini, Claudia Lefimil, Felipe A. Veloso, Inti Pedroso, David S. Holmes and Eugenia Jedlicki, "Bioinformatic prediction and experimental verification of Fur-regulated genes in the extreme acidophile Acidithiobacillus ferrooxidans", Nucleic Acids Research, Vol. 35, No. 7, pp. 2153-2166, 2007
- [35] Naveed Massjouni, Corban G. Rivera and Murali, "VIRGO: computational prediction of gene functions", Nucleic Acids Research, Vol. 34, No.2, pp. 340-344, 2006
- [36] Thomas E. Royce, Joel S. Rozowsky and Mark B. Gerstein, "Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification", Nucleic Acids Research, Vol.35, No.15, 2007
- [37] Xiaomei Wu, Lei Zhu, Jie Guo, Da-Yong Zhang and Kui Lin, "Prediction of yeast protein-protein interaction network: insights from the Gene Ontology and annotations", Nucleic Acids Research, Vol.34, No.7, pp.2137-2150, April 2006
- [38] Sung-Kyu Kim, Jin-Wu Nam, Je-Keun Rhee, Wha-Jin Lee and Byoung-Tak Zhang, "miTarget: microRNA target gene prediction using a support vector machine", BMC Bioinformatics, Vol.7, No.411, pp.1-14, 2006
- [39] Marijke J. van Baren and Michael R. Brent, "Iterative gene prediction and pseudogene removal improves genome annotation", Genome Research, Vol.16, pp.678-685, 2006
- [40] Richard A. George, Jason Y. Liu, Lina L. Feng, Robert J. Bryson-Richardson, Diane Fatkin and Merridee A. Wouters, "Analysis of protein sequence and interaction data for candidate disease gene prediction", Nucleic Acids Research, Vol.34, No.19, pp.1-10, 2006
- [41] Gustavo Glusman, Shizhen Qin, Raafat El-Gewely, Andrew F. Siegel, Jared C. Roach, Leroy Hood and Arian F. A. Smit, "A Third Approach to Gene Prediction Suggests Thousands of Additional Human Transcribed Regions", PLOS Computational Biology, Vol.2, No.3, pp.160-173, March 2006
- [42] Hongwei Wu, Zhengchang Su, Fenglou Mao, Victor Olman and Ying Xu, "Prediction of functional modules based on comparative genome analysis and Gene Ontology application", Nucleic Acids Research, Vol.33, No.9, pp.2822-2837, 2005
- [43] Yanhong Zhou, Huili Zhang, Lei Yang and Honghui Wan, "Improving the Prediction Accuracy of Gene structures in Eukaryotic DNA with Low C+G Contents", International Journal of Information Technology Vol.11, No.8, pp.17-25,2005
- [44] Reese, Kulp, Tammana, "Genie - Gene Finding in Drosophila Melanogaster", Genome Research, Vol.10, pp.529-538, 2000
- [45] Philippe P. Luedi, Alexander J. Hartemink and Randy L. Jirtle, "Genome-wide prediction of imprinted murine genes", Genome Research, Vol.15, pp. 875-884, 2005
- [46] Mohammed Zahir Hossain Sarker, Jubair Al Ansary and Mid Shajjad Hossain Khan, "A new approach to spliced Gene Prediction Algorithm", Asian Journal of Information Technology, Vol.5, No.5, pp.512-517, 2006
- [47] Said S. Adi and Carlos E. Ferreira, "Gene prediction by multiple syntenic alignment", Journal of Integrative Bioinformatics, Vol.2, No.1, 2005
- [48] Mario Stanke and Burkhard Morgenstern, "AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints", Nucleic Acids Research, Vol.33, pp.465-467, 2005
- [49] Kashiwabara, Vieira, Machado-Lima and Durham, "Splice site prediction using stochastic regular grammars", Genet. Mol. Res, Vol. 6, No.1, pp.105-115, 2007
- [50] Xiaobo Zhou, Xiaodong Wang and Edward R.Dougherty, "Gene Prediction Using Multinomial Probit Regression with Bayesian Gene Selection", EURASIP Journal on Applied Signal Processing, Vol.1, pp.115-124, 2004
- [51] Jonathan E. Allen, Mihaela Pertea and Steven L. Salzberg, "Computational Gene Prediction Using Multiple Sources of Evidence", Genome Research, Vol.14, pp.142-148, 2004
- [52] Biju Issac and Gajendra Pal Singh Raghava, "EGPred: Prediction of Eukaryotic Genes Using Ab Initio Methods after combining with sequence similarity approaches", Genome Research, Vol.14, pp.1756-1766, 2004
- [53] Leila Taher, Oliver Rinner, Saurabh Garg, Alexander Sczyrba and Burkhard Morgenstern, "AGeDA: gene prediction by cross-species sequence comparison", Nucleic Acids Research, Vol. 32, pp.305-308, 2004
- [54] Manpreet Singh, Parminder Kaur Wadhwa, and Surinder Kaur, "Predicting Protein Function using Decision Tree", World Academy of Science, Engineering and Technology, Vol.39, No. 66, pp.350-353, 2008
- [55] Trevor W. Fox and Alex Carreira, "A Digital Signal Processing Method for Gene Prediction with Improved Noise Suppression", EURASIP Journal on Applied Signal Processing, Vol.1, pp.108-114, 2004
- [56] Kai Wang, David Wayne Ussery and Søren Brunak, "Analysis and prediction of gene splice sites in four Aspergillus genomes", Fungal Genetics and Biology, Vol. 46, pp.14-18, 2009
- [57] Mai S. Mabrouk, Nahed H. Solouma, Abou-Bakr M. Youssef and Yasser M. Kadah, "Eukaryotic Gene Prediction by an Investigation of Nonlinear Dynamical Modeling Techniques on EIP Coded Sequences", International Journal of Biological and Life Sciences, Vol. 3, No.4, pp. 225-230, 2007
- [58] Yingyao Zhou, Jason A. Young, Andrey Santosyan, Kaisheng Chen, S. Frank Yan and Elizabeth A. Winzler, "In silico gene function prediction using ontology-based pattern identification", Bioinformatics, Vol.21, No.7, pp.1237-1245, 2005
- [59] Poonam Singhal, Jayaram, Surjit B. Dixit and David L. Beveridge, "Prokaryotic Gene Finding Based on Physicochemical Characteristics of Codons Calculated from Molecular Dynamics Simulations", Biophysical Journal, Vol.94, pp.4173-4183, June 2008
- [60] Thomas Schiex, Jerome Gouzy, Annick Moisan and Yannick de Oliveira, "FrameD: a flexible program for quality check and gene prediction in prokaryotic genomes and noisy matured eukaryotic sequences", Nucleic Acids Research, Vol.31, No.13, pp.3738-3741, 2003
- [61] ZHONG Yiming, JIANG Guanghui, CHEN Xuewei, XIA Zhihui, LI Xiaobing, ZHU Lihuang and ZHAI Wenxue, "Identification and gene prediction of a 24 kb region containing xa5, a recessive bacterial blight resistance gene in rice (Oryza sativa L.)", Chinese Science Bulletin, Vol. 48, No. 24, pp.2725-2729,2003
- [62] Gautam Aggarwal and Ramakrishna Ramaswamy, "Ab initio gene identification: prokaryote genome annotation with GeneScan and GLIMMER", J.Biosci, Vol.27, No.1, pp.7-14, February 2002
- [63] Shouguo Gao and Xujing Wang, "Predicting Type 1 Diabetes Candidate Genes using Human Protein-Protein Interaction Networks", J Comput Sci Syst Biol, Vol. 2, pp.133-146, 2009
- [64] Freudenberg and Propping, "A similarity-based method for genome-wide prediction of disease-relevant human genes", Bioinformatics, Vol. 18, No.2, pp.110-115, April 2002
- [65] Thao T. Tran, Fengfeng Zhou, Sarah Marshburn, Mark Stead3, Sidney R. Kushner and Ying Xu, "De novo computational prediction of non-coding RNA genes in prokaryotic genomes", Bioinformatics, Vol.25, No.22, pp.2897-2905, 2009

- [66] Pierre Montalent and Johann Joets, "EuGene-maize: a web site for maize gene prediction", *Bioinformatics*, Vol.26, No.9, pp.1254-1255, 2010
- [67] Zafer Barutcuoglu, Robert E. Schapire and Olga G. Troyanskaya, "Hierarchical multi-label prediction of gene functions", *Bioinformatics*, Vol.22, No.7, pp.830-836, 2006
- [68] Pernille Nielsen and Anders Krogh, "Large-scale prokaryotic gene prediction and comparison to genome annotation ", *Bioinformatics*, Vol.21, No.24, pp.4322-4329, 2005
- [69] Huiqing Liu, Jinyan Li and Limsoon Wong, "Use of extreme patient samples for outcome prediction from gene expression data", *Bioinformatics*, Vol.21, No.16, pp.3377-3384, 2005
- [70] Jiang Qian, Jimmy Lin, Nicholas M. Luscombe, Haiyuan Yu and Mark Gerstein, "Prediction of regulatory networks: genome-wide identification of transcription factor targets from gene expression data", *Bioinformatics*, Vol.19, No.15, pp.1917-1926, 2003
- [71] Shin Kawano, Kosuke Hashimoto, Takashi Miyama, Susumu Goto and Minoru Kanehisa, "Prediction of glycan structures from gene expression data based on glycosyltransferase reactions", *Bioinformatics*, Vol.21, No.21, pp.3976-3982, 2005
- [72] Alona Fyshe, Yifeng Liu, Duane Szafron, Russ Greiner and Paul Lu, "Improving subcellular localization prediction using text classification and the gene ontology", *bioinformatics*, Vol.24, No.21, pp.2512-2517, 2008
- [73] Jensen, Gupta, Starfeldt and Brunak, "Prediction of human protein function according to Gene Ontology categories", *Bioinformatics*, Vol.19, No.5, pp.635-642, 2003
- [74] Daniel Barker, Andrew Meade and Mark Pagel, "Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes", *Bioinformatics*, Vol.23, No.1, pp.14-20, 2007
- [75] Takatsugu Kan, Yutaka Shimada, Funiaki Sato, Tetsuo Ito, Kan Kondo, Go Watanabe, Masato Maeda, eiji Yamasaki, Stephen J.Meltzer and Masayuki Imamura, "Prediction of Lymph Node Metastasis with Use of Artificial Neural Networks Based on Gene Expression Profiles in Esophageal Squamous Cell Carcinoma", *Annals of surgical oncology*, Vol.11, No.12, pp.1070-1078, 2004
- [76] Shaun Mahony, Panayiotis V. Benos, Terry J.Smith and Aaron Golden, "Self-organizing neural networks to support the discovery of DNA-binding motifs", *Neural Networks*, Vol.19, pp.950-962, 2006
- [77] Zainal A. Hasibuan, Romi Fadhilah Rahmat, Muhammad Fermi Pasha and Rahmat Budiarto, "Adaptive Nested Neural Network based on human Gene Regulatory Network for gene knowledge discovery engine", *International Journal of Computer Science and Network Security*, Vol.9, No.6, pp.43-54, June 2009
- [78] Liu Qicai, Zeng Kai, Zhuang Zehao, Fu Lengxi, Ou Qishui and Luo Xiu, "The Use of Artificial Neural Networks in Analysis Cationic Trypsinogen Gene and Hepatitis B Surface Antigen", *American Journal of Immunology*, Vol.5, No.2, pp.50-55, 2009
- [79] Alistair M. Chalk and Erik L.L. Sonnhammer, "Computational antisense oligo prediction with a neural network model", *Bioinformatics*, Vol.18, No.12, pp.1567-1575, 2002

AUTHORS PROFILE



Manaswini Pradhan received the B.E. in Computer Science and Engineering, M.Tech in Computer Science from Utkal University, Orissa, India. She is into teaching field from 1998 to till date. Currently she is working as a Lecturer in P.G. Department of Information and Communication Technology, Orissa, India. She is currently persuing the Ph.D. degree in the P.G. Department of Information and communication

Technology, Fakir Mohan University, Orissa, India. Her research interest areas are neural networks, soft computing techniques, data mining, bioinformatics and computational biology.



Dr Ranjit Kumar Sahu, M.B.B.S, M.S. (General Surgery), M. Ch. (Plastic Surgery). Presently working as an Assistant Surgeon in post doctoral department of Plastic and reconstructive surgery, S.C.B. Medical College, Cuttack, Orissa, India. He has five years of research experience in the field of surgery and published one international paper in Plastic Surgery.

A Multicast Framework for the Multimedia Conferencing System (MCS) based on IPv6 Multicast Capability

¹ Hala A. Albaroodi ²Omar, Amer Abouabdalla ³Mohammed Faiz Aboalmaaly and ⁴Ahmed M. Manasrah

National Advanced IPv6 Centre
Universiti Sains Malaysia
Penang, Malaysia

Abstract- This paper introduces a new system model of enabling the Multimedia Conferencing System (MCS) to send a multicast traffic based on IPv6. Currently, the above mentioned system is using a unicast approach to distribute the multimedia elements in an IPv4-based network. Moreover, this study covers the proposed system architecture as well as the expected performance gain for transforming the current system from IPv4 to IPv6 by taking into account the advantages of IPv6 such as the multicast. Expected results shows that moving the current system to run on IPv6 will dramatically reduce the network traffic generated from IPv4-based MCS

Keywords- IPv6 Multicast, Multimedia Conference, MCS;

I. INTRODUCTION:

In the Last few years, the numbers of Internet users have increased significantly. Accordingly, the internet services are increased as well with taking into account their scalability and robustness. In terms of internet's transmission mode in IPv4, there are two types available, namely; unicast and multicast. As the name implies, unicast is a one to one communication, in other word, each packet will be transferring from one source to one destination, and while in contrast, the multicasting is the

the mechanism used to transmit the multimedia traffic among the users, section four outlines the proposed MCS which makes use of IPv6 multicasting for transmitting the multimedia content. We conclude our work in section 5 and we end our paper by the reference in section 6.

II. IPv6 MULTICASTING

IP multicasting is better than unicast in that it enables a source host to transfer a single packet to one or more destination hosts, which are recognized by a single group address. The packets are duplicated inside the IP network by routers, while only one packet, destined for a specific host, will be sent to a complete-link. This keeps bandwidth low at links leading to multiple destination

way of single packet will duplicate at the source's side or the router's side into many identical packets to reach many destinations. Additionally, in IPv4 special class used for multicasting which is a class D IP addressing and other classes are usually used for unicasting. We do not want to go in details with the unicasting since it is out of our scope of this study, but in the meantime we will focus only on the multicast approach.

In IPv4, multicasting has some cons in general because it is required a multicast routers and some other issues related to packet dropping. Moreover, in order to a wide adoption for a given software or application, the presence of infrastructure for that particular software or application is important, from this point we have seen that there is no "enough" IPv4 multicast infrastructure available today. Furthermore, most of the studies are now focusing on building application based on the IPv6 in general since it is the next generation of the IP.

The rest of this paper is organized as follow. In the next section, an overview to the IPv6 Multicasting is addressed, while in section three we introducing our MCS product as an audiovisual conferencing system and discusses its structure in terms of

hosts in the same group. A source host only has to know one group address to reach an arbitrarily sized group of destination hosts. IP multicasting is designed for applications and services in which the same data needs to concurrently reach many hosts joined in a network; these applications include videoconferencing, company communication, distance learning, and news broadcasting.

IP multicasting offers an alternative to normal unicast; in which the transferring source host can support these applications by learning the IP addresses of n destination hosts, establishing n point-to-point sessions with them, and transmitting n copies of each packet. Due to these characteristics, an IP multicast solution is more

effective than traditional broadcasting and is less of a resource burden on the source host and network

III. THE CURRENT MCS

The current MCS was introduced by [1] "A Control Criteria to Optimize Collaborative Document and Multimedia Conferencing Bandwidth Requirements". The current MCS implemented by the Network Research Group (NRG) from the school of Computer science at the University Science Malaysia in collaboration with Multimedia Research Labs Sdn. Bhd. The author defines current MCS that utilizes a switching method to obtain low bandwidth consumption, which until now allows an unlimited number of users to participate in the conference. He also describes a set of conference control options that can be considered as rules for controlling the current MCS and are called Real-time Switching (RSW) control criteria [2].

Today, most of the video conferencing systems available require high bandwidth and consume a large share of system resources. On the other hand, the current MCS design was based on a distributed architecture, which allows a form of distributed processing to support multimedia conferencing needs. In addition, this distributed design can be easily adapted to comply with any network structure [3].

The current MCS is one of the applications that use multicasting to achieve multipoint-to-multipoint conferencing. The MCS currently uses IPv4 multicasting only within a single Local Area Network (LAN). It uses Multiple LAN IP Converter (MLIC) to distribute audio and video through the WAN or Internet; this generates unnecessary packets, since MLIC uses unicasting technology to deliver these packets to current MCS conference participants located in different LANs. The MLIC will convert unicast packets to multicast only when delivering audio and video packets to conference participants located in the same LAN the MLIC connected to.

The current MCS has four main components (Current MCS Server, Current MCS Client, MLIC and Data compression / Decompression component). Each component has a task list and can be plugged into a network and unplugged without crashing the system. The current MCS server is the only component that will shut down the entire system if it is unplugged or shut down. The current MCS components are also called entities, and they have the ability to reside anywhere on the network, including sharing the same host as other network entities. Currently, the current MCS server and MLIC share one host while the current MCS client and the Data compression/decompression component share a different

host [4,5]. The general architecture of the current MCS components is shown in Figure 1.

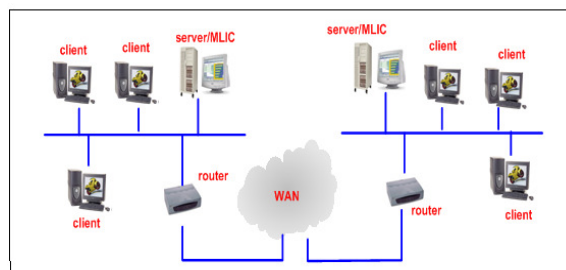


Figure 2.1: The Current MCS General Architecture.

IV. THE MLIC ENTITY

The MLIC is needed when more than one IP LAN is involved in the multimedia conference. This is because the UDP packets transmitted by the client object are IPv4 multicast packets. Most routers will drop IPv4 multicast packets since it is not recognized over Internet, and thus the multicast audio and video UDP packets will never cross over a router. The job of an MLIC is to function as a bi-directional tunnelling device that will encapsulate the multicast packets in order to transport them across routers, WANs and the Internet.

The MLIC has two interfaces: the LAN interface and the router interface. All MLICs are bi-directional and can provide reception and transmission at the same time. MLICs can also handle more than one conference at a time. The functions of the MLIC can be defined as follows:

- i. Audio/Video packets are transmitted by the client (active site) in LAN 1; MLIC in LAN 1 will do the following:
 - a. Listen on the specified port for Audio/Video UDP multicast packets.
 - b. Convert multicast packets to Audio/Video UDP unicast packets and transmit them.
- ii. The converted packets then go through the WAN router to LAN 2; the MLIC in LAN 2 will then:
 - a. Receive Audio/Video UDP unicast packets from the MLIC in LAN 1.
 - b. Convert Audio/Video UDP unicast to Audio/Video UDP multicast packets and retransmit within LAN 2.

Figure 2.2 shows the network architecture including MLICs.

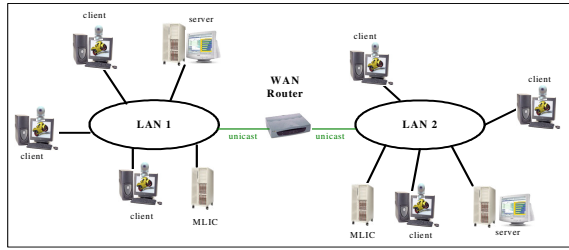


Figure 2: Multi LAN with MLICs

Additionally, several video conferencing systems are existed nowadays in the market. Each one of them has its own advantage as well as some disadvantage. The most important literature view limitation of this study can be summarized in Table 1. The limitation in the existing system can be addressed and overcome by using multicasting capability over IPv6 to deliver audio and video to the multimedia conferencing participants.

TABLE 1. PREVIOUS WORK'S LIMITATION

System	Explanation	Limitation
Current MCS	The current MCS is a conferencing system that allows clients to confer by using video and audio. Current MCS uses MLIC to distribute audio and video through the WAN or Internet.	MLIC is an application layer entity that cause delay in audio and video delivery. MLIC uses unicast, which generate unnecessary traffic.
HCP6 [6]	The HCP6 is a high quality conferencing platform. The audio is encoded in MP3 format. The HCP6 video is encoded in MPEG4 format.	It uses IPv6 multicast for audio and video delivery. Substantial end-to-end delay may be caused due to the double buffering that used to transfer audio and video. This is not suitable for interactive communications.
VIC [7]	VIC only provides the video part of the	It uses multicast over IPv4, which is usually

	multimedia conferencing. Audio, WB, and control tools are implemented separately.	dropped by routers.
HVCT [8]	This study is focused on designing and implementation a high-quality video conferencing tool based on IPv6 capability.	It uses the encoding / decoding operations and the built-in multiplexing / de-multiplexing operations that causes a delay. It is the main limitation of this study. Furthermore, delay is a very important factor especially in real-time applications.
VLC [9]	VLC media player is a portable multimedia player for various audio and video formats like MPEG-1, MPEG-2, MPEG-4, DivX, mp3. In addition to that VLC has the capability to plays DVDs, VCDs, and various formats. The system components are VLS (VideoLAN Server) and the VLC (VideoLAN Client).	There are still some features of the VLC media player which do not support IPv6. In particular, it is impossible to use RTSP over IPv6 because the underlying library, Live.com, does not support IPv6 at the time of writing. VLC by default uses IPv4.
VideoPort SBS Plus [10]	It uses IPv4 multicast capability to deliver the multimedia packets among the participant.	It uses multicast over IPv4, which is usually dropped by routers.

V. THE PROPOSED MCS

The new MCS system was composed to serve several different purposes. This program implements the new MCS system, consisting of clients and a server. Both client and server determine the type of the message, whether it be a request or a response to a request. A request message carries requests from the client to the server, while a response message carries responses from the server to the client.

When a client wants to start a conference using the program, the client is required to login to the server. Once the username and password are verified, the client will be able to create new conference or join existing conference. The client can select the participant with whom she/he wished to confer. After the other participants are selected, an invitation will be sent to the participants and the chairman will join a multicast group, which assigned by the server. Once the invitation is accepted, the clients can join the multicast group and can then begin voice or video conversation. Any client who is currently in a conversation group will not be available for another conversation group. Clients can log off from the conversation group by clicking on the "leave group" button, clients can logged off from the server, terminating any further possibility of conversation.

This study focuses on delivering audio and video using IPv6 multicast. This process will not only save time in capturing and converting the packet, but will also minimize bandwidth usage. The new globally recognizable multicast addresses in IPv6 allow new MCS multicast packets to be directly routed to the other clients in different LANs.

The New MCS Process

Multicasting helps to achieve this process, which depends on a set of rules permitting smooth flow from the creation of the conference to its termination. The steps involved in the process are listed below:

- i. Logging in.
- ii. Creating the conference.
- iii. Inviting participants to the conference.
- iv. Joining the conference.
- v. Transferring of Audio and Video.
- vi. Terminating the conference.

Network application requirements are developing rapidly, especially in audio and video applications. For this reason, these researches propose new MCS, which uses IPv6 multicasting to obtain speed and high efficiency without increasing bandwidth. This can be achieved without using MLIC. The proposed architecture provides

a complete solution to make multicast-based wide-area audio and video conferencing possible. The following steps, along with Figure 3.4, briefly illustrate how data transfer and user processes will occur in this new MCS.

- i. First, users should login to the server.
- ii. Users then can start a new conference or join an existing conference.
- iii. Clients will request the IPv6 multicast address from the server.
- iv. The server will assign unique multicast address to each conference.

The flowchart below shows the steps involved for starting a multimedia conference using the proposed MCS.

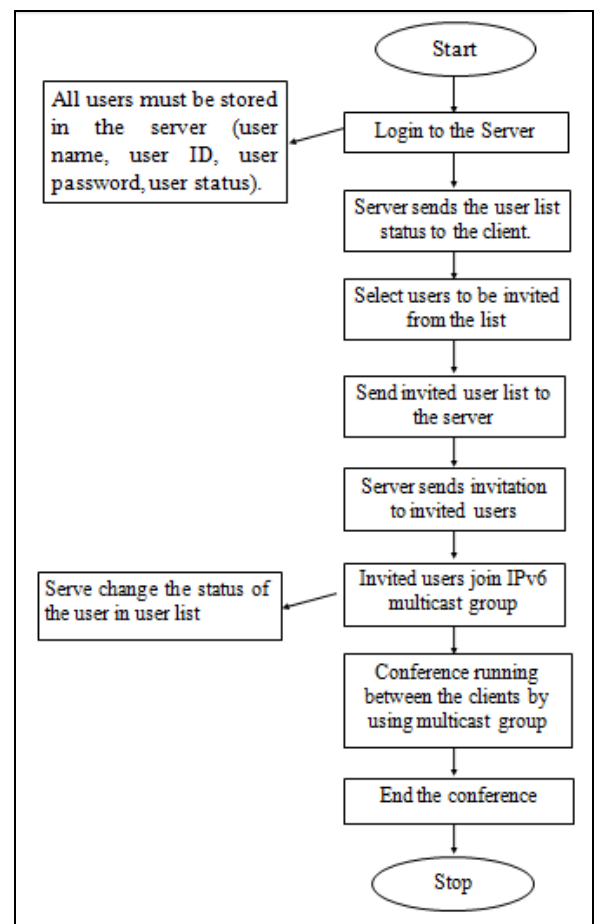


Figure 3 Steps of the New MCS

VI. CONCLUSION AND FUTURE WORKS

In this study, all the video and audio packets will be transmitted via an IPv6 multicasting. Due to the nature of multicasting, packets sent only once in client side. All participants will be able to receive the packets without

any issue. With this also, network congestions will be reduces drastically with the help of single multicast packet sending instead of multiple unicast packets.

The new MCS improve bandwidth consumption by using lower bandwidth than current MCS. With the new MCS, many organizations that have limited bandwidth will be able to use the implementation and obtain optimal results. Finally, the system developed in this research could also contribute to the reduction of network congestion when using multimedia conferencing system.

This work focused mainly on audio and video communication among MCS users by adopting the IPv6 multicasting capability. Current MCS is able to provide several services to the user and not only audio and video communication, such as application conferencing (AC) and document conferencing (DC), both features are currently working over IPv4. Since better network bandwidth utilization has been gaining from running the new module, Migration AC and DC to be worked over IPv6 will mostly reduce the overall utilized bandwidth by the current MCS application.

REFERENCES

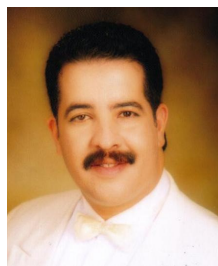
- [1] RAMADASS, S. (1994) A Control Criteria to Optimize Collaborative Document and Multimedia Conferencing Bandwidth Requirements. International Conference on Distributed Multimedia Systems and Applications (ISMM). Honolulu, Hawaii: ISMM. pp 555-559.
- [2] KOLHAR, M. S., BAY AN, A. F., WAN, T. C., ABOUABDALLA, O. & RAMADASS, S. (2008) Control and Media Sessions: IAX with RSW Control Criteria. International Conference on Network Applications, Protocols and Services 2008 (NetApps2008) Executive Development Center, Universiti Utara Malaysia. pp 75-79.
- [3] RAMADASS, S., WAN, T. C. & SARAVANAN, K. (1998) Implementing The MLIC (Multiple LAN IP Converter). Proceedings SEACOMM'98. pp 12-14.
- [4] BALAN SINNIH, G. R. S., & RAMADASS, S. (2003) Socket Level Implementation of MCS Conferencing System in IPv6 IN KAHNG, H.-K. (Ed.) International Conference, ICOIN 2003. Cheju Island, Korea, Springer, 2003. pp 460-472.
- [5] GOPINATH RAO, S., ETTIKAN KANDASAMY, K. & RAMADASS, S. (2000) Migration Issues of MCSv4 to MCSv6. Proceeding Internet Workshop 2000. Tsukuba, Japan pp 14-18.
- [6] YOU, T., MINKYO, I., SEUNGYUN, L., HOSIK, C., BYOUNGWOOK, L. & YANGHEE, C. (2004). HCP6: a high-quality conferencing platform based on IPv6 multicast. Proceedings of the 12th IEEE International Conference on Networks, 2004. (ICON 2004. pp 263- 267.
- [7] MCCANN, S. & JACOBSON, V. (1995) vic: A Flexible Framework for Packet Video Proceedings of the third ACM international conference on Multimedia. San Francisco, California, United States, ACM. pp 511-522.
- [8] YOU, T., CHO, H., CHOI, Y., IN, M., LEE, S. & KIM, H. (2003) Design and implementation of IPv6 multicast based High-quality Videoconference Tool (HVCT).
- [9] VLC, VideoLAN (2009) [Online] [31st May 2009]. Internet: <http://wiki.videolan.org/Documentation:Play_HowTo/Introduction_to_VLC>.
- [10] VideoPort SBS Plus (2009) [Online] [31st May 2009] Internet: <http://video-port.com/docs/VideoPort_SBS_Plus_eng.pdf>

AUTHORS PROFILE



Hala A. Albaroodi, A PhD candidate joined the NAV6 in 2010. She received her Bachelor degree in computer sciences from Mansour University College (IRAQ) in 2005 and a master's degree in computer sciences from Univeriti Sains Malaysia (Malaysia) in 2009. Her PhD research is on peer-to-peer computing. She has numerous

research of interest such as IPv6 multicasting and video Conferencing.



Dr. Omar Amer Abouabdalla obtained his PhD degree in Computer Sciences from University Science Malaysia (USM) in the year 2004. Presently he is working as a senior lecturer and domain head in the National Advanced IPv6 Centre – USM. He has published more than 50 research articles in Journals and Proceedings

(International and National). His current areas of research interest include Multimedia Network, Internet Protocol version 6 (IPv6), and Network Security.



Mohammed Faiz Aboalmaali, A PhD candidate, He received his bachelor degree in software engineering from Mansour University College (IRAQ) and a master's degree in computer science from Univeriti Sains Malaysia (Malaysia). His PhD. research is mainly focused on Overlay Networks. He is interested in

several areas of research such as Multimedia

Conferencing, Mobile Ad-hoc Network (MANET) and Parallel Computing.



Dr. Ahmed M. Manasrah is a senior lecturer and the deputy director for research and innovation of the National Advanced IPv6 Centre of Excellence (NAV6) in Universiti Sains Malaysia. He is also the head of inetmon project “network monitoring and security monitoring platform”. Dr. Ahmed

obtained his Bachelor of Computer Science from Mu’tah University, al Karak, Jordan in 2002. He obtained his Master of Computer Science and doctorate from Universiti Sains Malaysia in 2005 and 2009 respectively. Dr. Ahmed is heavily involved in researches carried by NAV6 centre, such as Network monitoring and Network Security monitoring with 3 Patents filed in Malaysia.

THE EVOLUTION OF CHIP MULTI-PROCESSORS AND ITS ROLE IN HIGH PERFORMANCE AND PARALLEL COMPUTING

A.Neela madheswari,
Research Scholar, Anna University, Coimbatore,
India.

Dr.R.S.D.Wahida banu,
Research Supervisor, Anna University, Coimbatore,
India.

Abstract - The importance given for today's computing environment is the support of a number of threads and functional units so that multiple processes can be done simultaneously. At the same time, the processors must not suffer from high heat liberation due over increase in frequencies to attain high speed of the processors and also they must attain high system performance. These situations led to the emergence and the growth of Chip Multi-Processor (CMP) architecture, which forms the basis for this paper. It gives the contribution towards the role of CMPs in parallel and high performance computing environments and the needs to move towards CMP architectures in the near future.

Keywords- CMPs; High Performance computing; Grid Computing; Parallel computing; Simultaneous multithreading.

I. INTRODUCTION

Advances in semiconductor technology enable the integration of billion transistors on a single chip. Such exponentially increasing transistor counts makes reliability an important design challenge since a processor's soft error rate grows in direct proportion to the number of devices being integrated [7]. The huge amount of transistors, on the other hand, leads to the popularity of multi-core processor or chip multi-processor architectures for improved system throughput [13].

Multi-core processors represents an evolutionary change in conventional computing as well setting the new trend for high performance computing (HPC) - but parallelism is nothing new. Intel has a long history with the concept of parallelism and the development of hardware-enhanced threading capabilities. Intel has been delivering threading capable products for more than a decade. The move towards chip-level multiprocessing architectures with a large

number of cores continues to offer dramatically increased performance and power characteristics [14].

In recent years, Chip Multi-Processing (CMP) architectures have been developed to enhance performance and power efficiency through the exploitation of both instruction-level and thread-level parallelism. For instance, the IBMPower5 processor enables two SMT threads to execute on each of its two cores and four chips to be interconnected to form an eight-core module [8]. Intel Montecito, Woodcrest, and AMDAMD64 processors all support dual-cores [9]. Sun also shipped eight-core 32-way Niagara processors in 2006 [10, 15]. Chip Multi-Processors (CMP) have the advantages of:

1. Parallelism of computation: Multiple processors on a chip can execute process threads concurrently.
2. Processor core density in systems: Highly scalable enterprise class servers systems as well as rack-mount servers can be built that fit in several processor cores in a small volume.
3. Short design cycle and quick time-to-market: Since CMP chips are based on existing processor cores the product schedules can be short [5].

II. MOTIVATION

For the last few years, the software industry has significant advances in computing and the emerging grid computing, cloud computing and Rich Internet Applications will be the best examples for distributed applications. Although we are in machine-based computing now, a shift towards human-based computing are also emerging in which the voice, speech, gesture and commands of the human can be understand by the computers and act according to the human signals. Video conferencing, natural language processing and speech recognition software are come under this human-based computing as example. For these kinds of computing, there is a need for huge computing power with a number

of processors together with the advancement in multi-processor technologies.

In this decade, computer architecture has entered a new 'multi-core' era with the advent of Chip Multi-processors (CMP). Many leading companies, Intel, AMD and IBM, have successfully released their multi-core processor series, such as Intel IXP network processors [28], the Cell processor [12], the AMD Opteron™ etc. CMPs have evolved largely due to the increased power consumption in nanoscale technologies which have forced the designers to seek alternative measures instead of device scaling to improve performance. Increasing parallelism with multiple cores is an effective strategy [18].

III. EVOLUTION OF PROCESSOR ARCHITECTURE

Dual and multi-core processor systems are going to change the dynamics of the market and enable new innovative designs delivering high performance with an optimized power characteristic. They drive multithreading and parallelism at a higher than instruction level, and provide it to mainstream computing on a massive scale. From an operating system level (OS), they look like a symmetric multi-processor system (SMP) but they bring lot more advantage than typical dual or multi-processor systems.

Multi-core processing is a long-term strategy for Intel that began more than a decade ago. Intel has more than 15 multi-core processor projects underway and it is on the fast track to deliver multi-core processors in high volume across off of their platform families. Intel's multi-core architecture will possibly feature dozens or even hundreds of processor cores on a single die. In addition to general-purpose cores, Intel multi-core processors will eventually include specialized cores for processing graphics, speech recognition algorithms, communication protocols, and more. Many new and significant innovations designed to optimize the power, performance, and scalability is implemented into the new multi-core processors [14].

According to the number of functional units running simultaneously, the processor architecture is classified into 3 main types namely:

- (1) Single processor architecture, which does not support multiple functional units to run simultaneously.
- (2) Simultaneous multithreading (SMT) architecture, which supports multiple threads to run simultaneously but not the multiple functional units at any particular time.
- (3) Multi-core architecture or Chip multi-processor (CMP) architecture, which supports functional units to run simultaneous and may support multiple threads also simultaneously at any particular time.

A. Single processor architecture

The single processor architecture is shown in figure 1. Here only one processing unit is present in the chip for performing the arithmetic or logical operations. At any particular time, only one operation can be performed.

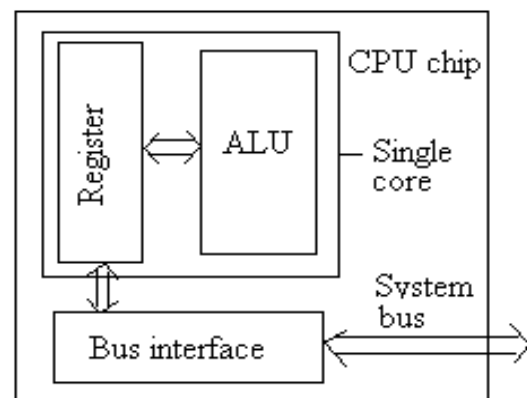


Figure 1: Single core CPU chip

B. Simultaneous multithreading (SMT) architecture

SMT permits simultaneous multiple independent threads to execute simultaneously on the same core. If one thread is waiting for a floating point operation to complete, another thread can use integer units. Without SMT, only a single thread can run at any given time. But in SMT, the same functional unit cannot be executed simultaneously. If two threads want to execute the integer unit at the same time, it is not possible with SMT. Here all the caches of the system are shared.

C. Chip Multi-Processor architecture

In multi-core or chip multi-processor architecture, multiple processing units or chips are present on a single die. Figure 2 shows a multi-core architecture with 3 cores in a single CPU chip. Here all the cores are fit on a single processor socket called as Chip Multi Processor. The cores can run in parallel. Within each core, threads can be time-sliced similar to single processor system [17].

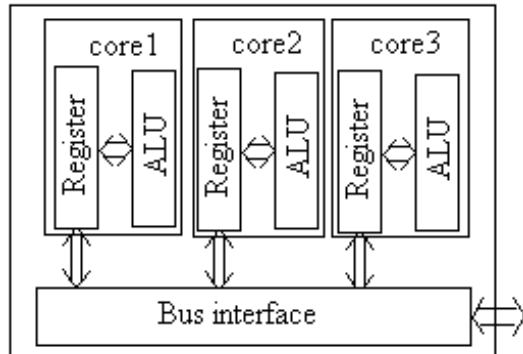


Figure 2: Chip multi-processor architecture

The multi-core architecture with cache and main memory is shown in Figure 3, comprises processor cores from 0 to N and each core has private L1 cache which consists of instruction cache (I-cache) and data cache (D-cache).

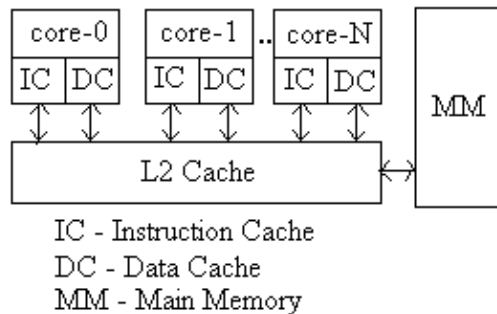


Figure 3: Multi-core architecture with memory

Each L1 cache is connected to the shared L2 cache. The L2 cache is unified and inclusive, i.e. it includes all the lines contained in the L1 caches. The main memory is connected to L2 cache, if the data requests are missed in L2 cache, the data access will happen in main memory [20].

IV. EXISTING ENVIRONMENTS FOR CHIP MULTI- PROCESSOR ARCHITECTURE

The chip multi-processors are used in the range of desktop to high performance computing environments. The section 4.1 and section 4.2 will show the existence and the main role of CMPs in various computing environments.

A. High Performance Computing

High performance computing uses super computers and computer clusters to solve advanced computation problems. A list of the most powerful high-performance computers can be found on the Top500 list.

Top500 is a list of the world's fastest computers. The list is created twice a year and includes some rather large systems. Not all Top500 systems are clusters, but many of them are built from the same technology. There may be HPC systems out there that are proprietary or not interested in the Top500 ranking. The Top500 list is the wealth of historical data. The list was started in 1993 and has data on vendors, organizations, processors, memory, and so on for each entry in the list [22]. As per the information taken at June 2010 from [23], the first 10 systems are given in the table 1.

Table 1: Top 10 Super computers list

Rank	Processor details	Year
1.	Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz.	2009.
2.	Nebulae - Dawning TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU.	2010.
3.	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 GHz / Opteron DC 1.8 GHz, Voltaire Infiniband.	2009.
4.	Kraken XT5 - Cray XT5-HE Opteron Six Core 2.6 GHz.	2009.
5.	JUGENE - Blue Gene/P Solution.	2009.
6.	Pleiades - SGI Altix ICE 8200EX/8400EX, Xeon HT QC 3.0/Xeon	2010.

	Westmere 2.93 GHz, Infiniband.	
7.	Tianhe-1 - NUDT TH-1 Cluster, Xeon E5540/E5450, ATI Radeon HD 4870 2, Infiniband.	2009.
8.	BlueGene/L - eServer Blue Gene Solution.	2007.
9.	Intrepid - Blue Gene/P Solution.	2007.
10.	Red Sky - Sun Blade x6275, Xeon X55xx 2.93 GHz, Infiniband.	2010.

Among the top 10 super computers, Jaguar and Kraken are having multi-core that are coming under CMP processors. Thus under high performance computing environments, the chip multi processors are involved and extends their capability in near future since the worldwide HPC market is growing rapidly. Successful HPC applications span many industrial, government and academic sectors.

B. Grid computing

Grid computing has emerged as the next-generation parallel and distributed computing methodology, which aggregates dispersed heterogeneous resources for solving various kinds of large-scale parallel applications in science, engineering and commerce [3]. As per [24], the list of the various grid computing environments are:

1. DAS-2: DAS-2 is a wide-area distributed computer of 200 Dual Pentium-III nodes [26].
2. Grid5000: It is distributed over 9 sites and contains approximately 1500 nodes and approximately 5500 CPUs [29].
3. NorduGrid: It is one of the largest production grids in the world having more than 30 sites of heterogeneous clusters. Some of the cluster nodes contain dual Pentium III processors [ng].
4. AuverGrid: It is a heterogeneous cluster [30].
5. Sharcnet: It is a cluster of clusters. It consists of 10 sites and has 6828 processors [24].
6. LCG: It contains 24115 processors [24].

Here the processors involved belong to multi core types under some grids. Hence under grid computing environment also chip multi-processors are used.

C. Parallel computing

Parallel computing plays a major role in the current trends and in almost all the fields. Formerly they are useful only to solve very huge problems such as weather forecasting, etc. But nowadays the concept of parallel computing are used starting from super computing environment to the modern desktop environment such as quad-core or in the GPU usage [25].

As per the parallel workload archive [21], the parallel computing systems are listed as:

1. CTC IBM SP2: It contains 512 nodes IBM SP2 during 1996.
2. DAS-2 5-Cluster: It contains 72 nodes, each of dual 1GHz Pentium-III during 2003.
3. HPC2N: It contains 120-node, each node contains two 240 AMD Athlon MP2000+ processors during 2002.
4. KTH IBM SP2: It contains 100 nodes IBM SP2 during 1996.
5. LANL: It contains 1024-node Connection Machine CM-5, during 1994.
6. LANL O2K: It contains a cluster of 16 Origin 2000 machines with 128 processors each (2048 total) during 1999.
7. LCG: It contains LHC (Large Hadron Collider) Computing Grid during 2005.
8. LLNL Atlas: It contains 1152 node, each node contains 8 AMD Opteron processors during 2006.
9. LLNL T3D: It contains 128 nodes, each node has two DEC Alpha 21064 processors. Each of the 128 nodes has two DEC Alpha 21064 processors during 1996.
10. LLNL Thunder: It contains 1024 nodes, each with 4 Intel IA-64 Itanium processors during 2007.
11. LLNL uBGL: It contains 2048 processors during 2006.
12. LPC: It contains 70 dual 3GHz Pentium-IV Xeon nodes during 2004.
13. NASA: It contains 128-nodes during 1993.

14. OSC Cluster: It has two types of nodes: 32 quad-processor nodes, and 25 dual-processor nodes, for a total of 178 processors during 2000.
15. SDSC: It contains 416 nodes during 1995.
16. SDSC DataStar: It contains 184 nodes during 2004.
17. SDSC Blue Horizon: It contains 144 nodes during 2000.
18. SDSC SP2: It contains 128-node IBM SP2 during 1998.
19. SHARCNET: It contains 10 clusters with quad and dual core processors during 2005.

Hence most of the processors involved in the parallel computing machines are multi-core processor types. This implies the involvement of multi-core processors in parallel computing environments.

V. CMP CHALLENGES

The advent of multi-core processors and the emergence of new parallel applications that take advantage of such processors pose difficult challenges to designers.

With relatively constant die sizes, limited on chip cache, and scarce pin bandwidth, more cores on chip reduces the amount of available cache and bus bandwidth per core, therefore exacerbating the memory wall problem [1]. The designer has to build a processor that provides a core with good single-thread performance in the presence of long latency cache misses, while enabling as many of these cores to be placed on the same die for high throughput.

Limited on chip cache area, reduced cache capacity per core, and the increase in application cache foot prints as applications scale up with the number of cores, will make cache miss stalls more problematic [19].

The problem of shared L2 cache allocation is critical to the effective utilization of multi-core processors. Sometimes unbalanced cache allocation will happen, and this situation can easily lead to serious problems such as thread starvation and priority inversion, which threatens to processor's utilization ratio and system performance.

Chip-Multiprocessor (CMP) or multi-core technology has become the mainstream in CPU designs. It embeds multiple processor cores into a single die to exploit thread-level parallelism for achieving higher overall chip-level Instruction-Per-Cycle (IPC) [2, 4, 6, 11, 27]. Combined with increased clock frequency, a multi-core, multithreaded processor chip demands higher on- and off-chip memory bandwidth and suffers longer average memory access delays despite an increasing on-chip cache size. Tremendous pressures are put on memory hierarchy systems to supply the needed instructions and data timely [16].

The memory and the chip memory bandwidth are a few of the main concern which plays an important role in improving the system performance in CMP architecture. Similarly the interconnection of the chips within the single die is also an important consideration.

VI. CONCLUSION

In today's scenario, it is essential to have a shift towards Chip multi processor architectures. It is not only applicable for the high performance and parallel computing but also for the desktops to face the challenges of system performance. Day by day, the challenges faced by the CMPs become complicated but the application and needs are also increasing. Suitable steps to be taken to decrease power consumption and leakage current.

References

- [1] W. Wulf and S. McKee, "Hitting the Memory Wall: Implications of the Obvious", ACM SIGArch Computer Architecture News, 23(1):20-24, March 1995.
- [2] L. Hammond, B. A. Nayfeh and K. Olukotun, A Single-Chip Multiprocessor, IEEE Computer, Sep. 1997.
- [3] I. Foster, C. Kesselman (Eds.), "The Grid: Blueprint for a Future Computing Infrastructure", Morgan Kaufmann Publishers, 1999.
- [4] J. M. Tendler, S. Dodson, S. Fields, H. Le, and B. Sinharoy, "IBM eserver Power4 System Microarchitecture," IBM White Paper, Oct. 2001.
- [5] Ishwar Parulkar, Thomas Ziaja, Rajesh Pendurkar, Anand D'Souza and Amitava Majumdar, "A Scalable, Low Cost Design-For-Test Architecture for UltraSPARC™ Chip Multi-

- Processors”, International Test Conference, IEEE, 2002, pp.726-735.
- [6] Sun Microsystems, “Sun’s 64-bit Gemini Chip,” Sunflash, 66(4), Aug. 2003.
- [ng] “Nordugrid – The Nordic Testbed for Wide area computing and data handling”, Final Report, Jan 2003.
- [7] S. Mukherjee, J. Emer, and S. Reinhardt, “The soft error problem, an architectural perspective”, HPCA-11, 2005.
- [8] B. Sinharoy, R. Kalla, J. Tendler, R. Eickemeyer, and J. Joyner. Power5 system microarchitecture. IBM Journal of Research and Development, 49(4/5):505–521, 2005.
- [9] C. McNairy and R. Bhatia. Montecito: A dual-core, dualthread itanium processor. IEEE Micro, 25(2):10–20, 2005.
- [10] P. Kongetira, K. Aingaran, and K. Olukotun. Niagara: A 32-way multithreaded sparc processor. IEEE Micro, 25(2):21–29, 2005.
- [11] AMD, Multi-core Processors: The Next Evolution in Computing, http://multicore.amd.com/WhitePapers/Multi-Core_Processors_WhitePaper.pdf, 2005.
- [12] A. Eichenberger, J. O’Brien, and et al. Using advanced compiler technology to exploit the performance of the cell broadband engine™ architecture. IBM Systems Journal, 45:59–84, 2006.
- [13] Huiyang Zhou, “A Case for fault tolerance and performance enhancement using Chip Multi-Processors”, IEEE Computer architecture letters”, Vol.5, 2006.
- [14] Pawel Gepner, Michal F.Kowalik, “Multi-Core Processors: New way to achieve high system performance”, In the proceedings of the International Symposium on Parallel computing in Electrical Engineering, IEEE, 2006.
- [15] Fengguang Song, Shirley Moore, Jack Dongarra, “L2 Cache Modeling for Scientific Applications on Chip Multi-Processors”, International Conference on Parallel Processing (ICPP), 2007.
- [16] Lu Peng, Jih-Kwon Peir, Tribuvan K. Prakash, Yen-Kuang Chen and David Koppelman, “Memory performance and scalability of Intel’s and AMD’s Dual-Core Processors: A case study”, IEEE, 2007, pp.55-64.
- [17] Jernej Barbic, “Multi-core architectures”, 15-213, Spring 2007, May 2007.
- [18] Sushu Zhang, Karam S.Chatha, “Automated Techniques for Energy Efficient scheduling on Homogeneous and Heterogeneous Chip Multi-processor architectures”, IEEE, 2008, pp.61-66.
- [19] Satyanarayana Nekkalapu, Haitham Akkary, Komal Jothi, Renjith Retnamma, Xiaoyu Song, “A Simple Latency Tolerant Processor”, IEEE, 2008, pp.384-389.
- [20] Benhai Zhou , Jianzhong Qiao, Shu-kuan Lin, “Research on fine-grain cache assignment scheduling algorithm for multi-core processors”, IEEE, 2009, pp.1-4.
- [21] Parallel workloads archive, Dror.G.Feitelson, <http://www.cs.huji.ac.il/labs/parallel/workload>, March 2009.
- [22] Douglas Eadline, “High Performance Computing for Dummies”, SUN and AMD Special edition, 2009.
- [23] Top 10 super computers, <http://www.top500.org/>, Sep 2010.
- [24] Grid computing environments, <http://gwa.ewi.tudelft.nl/pmwiki/>, June 2010.
- [25] A.Neela madheswari, R.S.D.Wahida banu, “Important essence of co-scheduling for parallel job scheduling”, Advances in Computational Sciences and Technology, Vol.3, No.1, 2010, pp.49-55.
- [26] The Distributed ASCI Supercomputer 2, <http://www.cs.vu.nl/das2/>, Sep 2010.
- [27] Intel, Inside Intel Core Microarchitecture and Smart Memory Access. <http://download.intel.com/technology/architecture/sma.pdf>.
- [28] Intel. Intel ixp2855 network processor - product brief.
- [29] Pierre Riteau, Mauricio Tsugawa, Andrea Matsunaga, Jose Fortes, Tim Freeman, Kate Keahey, “Sky computing on FutureGrid and Grid5000”.
- [30] AuverGrid, <http://gstatprod.cern.ch/gstat/site/AUVERGRID/>, Sep 2010.

AUTHOR’S PROFILE

A.Neela Madheswari received her Master of Computer Science and Engineering degree from Vinayaka Missions University, on June 2006. Currently, she is doing his research in the area of Parallel and Distributed systems under Anna University, Coimbatore. Earlier she completed her B.E, from Madras University of Computer Science and Engineering, Chennai on April 2000. Later, she joined as Lecturer at Mahendra Engineering College in CSE department from 2002. She had completed her M.E., from Vinayaka Missions University of Computer Science and Engineering during 2006 and now she serves as Assistant Professor at MET’S School of Engineering, Thrissur. Her research

interest includes Parallel and Distributed Computing and Web Technologies. She is a member of the Computer Society of India, Salem. She had presented the papers under national and international journals, national and international conferences. She is the reviewer in journals namely IJCNS and IJCSIS.

Towards a More Mobile KMS

Julius Olatunji Okesola

Dept. of Computer and Information
Sciences
Tai Solarin University of Education,
Ijebu-Ode, Nigeria

Oluwafemi Shawn Ogunseye

Dept. of Computer Science
University of Agriculture
Abeokuta, Nigeria

Kazeem Idowu Rufai

Dept. of Computer and Information
Sciences
Tai Solarin University of Education,
Ijebu-Ode, Nigeria

Abstract—Present knowledge management systems (KMS) hardly leverage the advances in technology in their designs. The effect of this cannot be positive because it creates avenues for dissipation and leaks in the knowledge acquisition and dissemination cycle. In this work we propose a development model that looks at KMS from the mobility angle enhancing previous designs of mobile KMS (mKMS) and KMS. We used a SOA based Smart Client Architecture to provide a new view of KMS with capabilities to actually manage knowledge. The model was implemented and tested as a small scale prototype to show its practicability. This model will serve as a framework and a guide for future designs.

Keywords- Knowledge Management; Service Oriented Architecture; Smart Client; Mobile KMS; Architecture
Introduction (Heading 1)

I. INTRODUCTION

Knowledge still remains the key resource for many organizations of the world. This is going to be the status quo for a long while. Organizations therefore attach a high level of importance to knowledge acquisition and dissemination. The understanding of this fact is however not fully appreciated nor obvious in the design of many KMSs. Tacit knowledge which is the major

source of competitive edge can be very transient, organizations that have the utmost value for knowledge would therefore understand the need for a system that can help acquire knowledge from experts or knowledge sources regardless of location and time and can also help disseminate knowledge to where it is needed when it is needed. We emphasize two concerns for consideration, firstly, Knowledge is only useful when it is applied [awad], but knowledge can only be applied when it is available when and where needed. This therefore requires KMS designs geared towards mobility. Secondly, since tacit knowledge can be generated in any instance, we need KMS's that is optimized to be available at those instances to facilitate acquisition of such knowledge for solving an organization's problems. These are issues that tend to emphasize a need for a more mobile oriented based design for KMSs. Mobility as referred to in this work goes beyond the use of mobile devices like Smart Phones, PDA's and mobile phones to access KMS, We instead proffer a model using current Service Oriented Architecture (SOA) and smart client architecture that can cut across different hardware platforms and

positions KMS for quick dissemination and acquisition of knowledge and other knowledge management functions to the benefit of the implementing organization. We do not limit our design to mobile devices like the previous reference models because of the fast disappearing line between the capabilities of mobile devices and computers. However, like the previous reference model, we however take into considerations the limitations of mobile devices [4], the limitations of organizations as regards location of experts and the individual limitation of the experts which can include, distractions, time pressure, work overload etc. We therefore build on previous research closing the gap between them and current possibilities and shed light on a potential way forward.

II. A NEW MODEL

Current KMS and mKMS design is really too network dependent helping only to retrieve and present knowledge resources to staff that are not within company premises but have access to company network [1& 2]. Our proposition improves on this by considering more than retrieval and presentation to acquisition and scalability. We also consider a bypass to intermittent connections through the design in such a way that if the staff is outside the reach of organization network for any reason, when he/she is within the network, they are immediately kept at par with any modifications or changes to sections of the knowledge base that affect them. They can also store knowledge on the device's light database/memory for upload to the

server at a later time. previous work [5], shows that the basic expectation of a mKMS are.

- facilitating the registration and sharing of insights without pushing the technique into the foreground and distracting mobile workers from the actual work,
- exploiting available and accessible resources for optimized task handling, whether they are remote (at home, in the office, or on the Web) or local (accompanying or at the customer's site), and
- privacy-aware situational support for mobile workers, especially when confronted with ad-hoc situations.

That is, mKM systems must not only provide mobile access to existing KM systems but also contribute to at least some of the above management goals.

A. SOA & Smart Clients

Service Oriented Architecture is an architectural paradigm that helps build infrastructure enabling those with needs (consumers) and those with capabilities (providers) to interact via services across disparate domain of technology and ownership [7]. SOA can enable the knowledge capabilities created by someone or a group of people be accessible to others regardless of where the creator(s) or consumer(s) is/are. It provides a powerful framework for matching needs and capabilities and for combining capabilities to address needs by leveraging other capabilities[7].

Smart clients can combine the benefits of rich client applications with the manageability and deployment of thin client applications

Combining SOA and Smart Clients provides the following capabilities[3]:

- Make use of local resources on hardware
- Make use of network resources

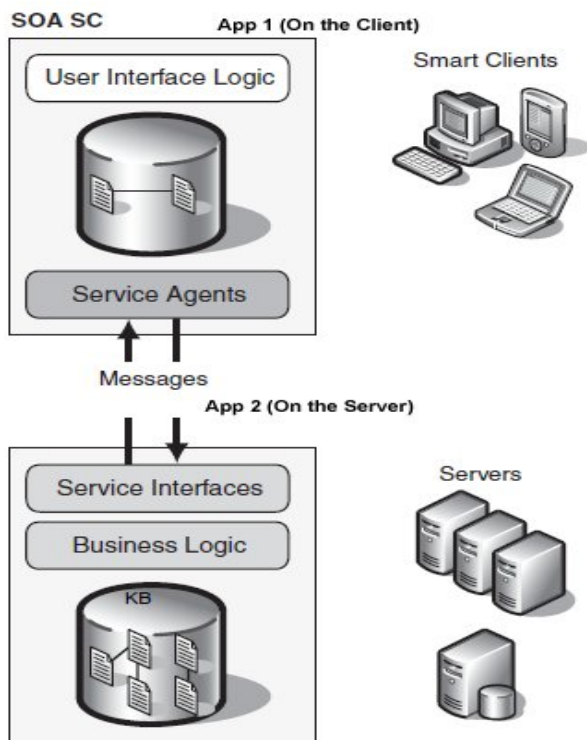
- Support occasionally connected users and field workers
 - Provide intelligent installation and update
 - Provide client device flexibility

These features are considered major advantages in improving KMS reach.

III. THE DESIGN

We propose a SOA based smart client model. The model can work with most mobile/computing device [3 & 6] and is not restricted to those that can use a database system. It also allows for loose coupling. The system's main business logic and data layer is situated on the server and a minor logic and application/presentation layer will reside on the user's machine.

Figure 1 below shows the overall architecture of our proposed model



The system will therefore have two parts, the server side application (App 2) & the client application (App 1).

1) At the client (App 1)

the system uses a thick client that can run on a wide range of devices from mobile devices to laptops. The smart client has the security information (login) and the user can use it to enter knowledge as it is generated in their field operations. The knowledge is synchronized with the company's knowledge base once they are within the reach of a network or onsite.

With App 1, the user will be able to store tacit knowledge as they are generated in the field. These knowledge which would normally be either scribbled down in jotters/ pieces of papers or forgotten (lost), can be saved and uploaded into the company's server when the user is within the reach of company network.

2) At the Server (App 2)

The server application comprises of a summarizer module. The module provides summary for knowledge solution which it sends to the client app/remote device. We employ on site synchronization between mobile device/computer with the KMS server. On site users can get the un-summarized version of the solution while the off shore users have to request. Further illustration is done through our sample application in the next section. The advantages of the new model are:

- decouples the client and server to allow independent versioning and deployment.
- reduces processing need on the client to a bearable minimum.
- gives more control and flexibility over data reconciliation issues.
- affords a lightweight client footprint.
- structures the KMS application into a service-oriented architecture.
- gives control over the schema of data stored on the client and flexibility that might be different from the server.

- The client application can interact with multiple or disparate services (for example, multiple Web services or services through Message Queuing, Web services, or RPC mechanisms).
- custom security scheme can be created.
- Allows the KMS application operate in an Internet or extranet environment.

many smart client applications are not able to support full relational database instances on the client. In such cases, the service-oriented approach shines bright ensuring that the appropriate infrastructure is in place to handle data caching and conflict resolutions [3].

The figure bellow depicts the message exchange pattern between the proposed model.

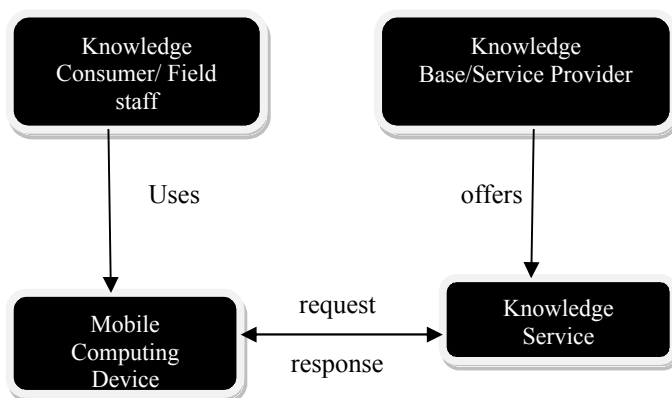


Figure 2: The interactions within the system

IV. APPLICATION

A prototype inter-Agency Criminal Knowledge and Intelligence System called the “Field Officer Knowledge Engine” (FOKE) was designed. The working of the system is described herein.

The FOKE prototype was designed to run on windows mobile 5.0 series customized for the specific purpose of running the FOKE. The aim of this prototype is to provide a platform for collaborative crime fighting between the

different government security agencies in the country. Since they are many agencies that fight specific crimes, they can have a central collaborative server to which criminals can be upgraded to based on certain criterion. Field agents of all agencies can be updated on current threats and criminals to watch out for regardless of where they are and they can share valuable findings with their collaborative communities of practice whenever the need arises without necessarily affecting their everyday task and individual goals.

The Full application resides on a development server for the purpose of testing, a lap top pc (serving as a regular client), the systems running the mobile device simulator and the development server are both allowed to connect to each other through wifi (ad hoc network). The simulation smart client was able to consume the services exposed by the application residing on the server when in the range of the wifi and when out of reach it cached data on the mobile device and laptop which it synchronized with the Knowledge base when connection was restored. The result of this simple implementation is shown in figures 3 and 4 below.



Figure 3: Login Page of the FOKE Prototype

The system is installed with user information locally stored. The system uses an Application Block {code} to detect availability of service indicated by the green label in figure 3.. The system detects the location of the officer when the officer is within range and requests password for authentication. Local data storage utilized a combination of long-term and short term data caching technique [3]. For the sake of security, the user PIN is store as in short-term caching so as to ensure volatility. Knowledge entered into the system by user is however stored through long term caching. When the user accesses a knowledge resources from the remote knowledge base server, the resource is stored through

a short –term caching to provide only quick revisits and save limited memory in mobile devices.



Figure 4: the Activity Page for the FOKE system

The application page served as the main presentation page. The system allowed for search through a search box, information/results returned were however highly filtered and summarized to avoid memory overload.

V. CONCLUSION & FUTURE WORK

The work showed how smart client and SOA can be combined to help extend the reach of KM practices through a proactive knowledge retrieval and knowledge acquisition model, the prototype implementation does not only shed light on 'how' it can be used to solve KM problems but also on where it can be used. The fact is smart client might be a little more restrictive than a thin client based model, because it implies that only specific kind of hardware can use it. This is an advantage for security.

From the sample implementation, It was seen that the design is indeed practicable and can serve as a framework for future design models of KMS. We did not give too much consideration to the issue of security in this model relying on basic security features of the system. This can enjoy more research and improvement.

References

- [1]. Awad E.M and Ghaziri H.M. (2004), Knowledge Management, Pearson Education Inc. New Jersey. 1st Ed.
- [2]. Matthias Grimm, Mohammad-Reza Tazari, and Dirk Balfanz (2005), A Reference Model for Mobile Knowledge Management, *Proceedings of I-KNOW '05 Graz, Austria, June 29 - July 1, 2005*
- [3]. David Hill, Brenton Webster, Edward A. Jezierski, Srinath Vasireddy, Mo Al-Sabt, Blaine Wastell, Jonathan Rasmusson, Paul Gale & Paul Slater(2004), Smart Client Architecture and Design Guide :patterns & practices. Microsoft Press
- [4]. Mobile Commerce: opportunities and challenges, A GS1 Mobile Com White Paper February 2008 Edition

[5]. Tazari, M.-R., Windlinger, L., and Hoffmann, T. (2005). Knowledge management requirements of mobile work on information technology. In *Mobile Work Employs IT (MoWeIT'05)*, Prague.

[6]. Mustafa Adaçal and Ayşe B. Bener, (2006), Mobile Web Services: A New Agent-Based Framework, IEEE INTERNET COMPUTING, pp 58-65

[7]. Nickull D., Reitman L., Ward J., and Wilber J. (2007), "Service Oriented Architecture (SOA) and specialized Messaging Patterns ". Adobe Systems Incorporated USA.

AUTHORS PROFILE

Dr. Julius O. Okesola is a Lecturer at the Department of Computer and Information Sciences, Tai Solarin University of Education, Ijebu-Ode, Ogun State, Nigeria. His areas of interest are: Information Systems, Multimedia Databases, Visualization, Computer Security, Artificial Intelligence & Knowledge Management.

Oluwafemi Shawn Ogunseye received his first degree in Computer Science from the University of Agriculture Abeokuta, Ogun State, Nigeria. He is an avid researcher. His areas of interest are: Information Systems, Computer & Information Security, Machine Learning & Knowledge Management.

Kazeem Idowu Rufai is a lecturer at the Tai Solarin University of Education, Ijebu-Ode Ogun State in Nigeria. He is an avid researcher whose research interest include Knowledge Management Systems, Computer Hardware Technology etc.

An Efficient Decision Algorithm for Vertical Handoff Across 4G Heterogeneous Wireless Networks

S.Aghalya
Research Scholar,
Anna University
India

P. Seethalakshmi
Anna University
India

Abstract - As mobile wireless networks increase in popularity, we are facing the challenge of integration of diverse wireless networks. It is becoming more important to arrive at a vertical handoff solution where users can move among various types of networks efficiently and seamlessly. To address this issue, an efficient vertical handoff decision (EVHD) algorithm has been proposed in this paper to decide the best network interface and best time moment to handoff. An overall gain function has been utilized in this algorithm to make the right decision based on various factors, the network characteristics such as usage cost, bandwidth, power consumption and dynamic factors such as Received Signal Strength (RSS), velocity and position of mobile terminal (MT). The effectiveness of the EVHD algorithm has been verified by carrying out simulations. The results show that EVHD achieves 78.5% reduction in number of unnecessary handoffs compared to static parameter based algorithm. The increase in throughput is about 60% compared to static parameter based algorithm for all the types of traffic. The overall system performance has been improved by the proposed efficient VHD algorithm and outperforms the three well known VHD algorithms including static parameter based, RSS based and RSS-timer based algorithms.

Keywords - *Heterogeneous network, Seamless handoff, Vertical handoff, Handoff decision, Gain function.*

I. INTRODUCTION

Nowadays, there are various wireless communication systems existing for different services, users and data rates such as GSM, GPRS, IS-95, W-CDMA, Wireless LAN etc. Fourth generation (4G) wireless systems integrate all existing and newly developed wireless access systems. 4G wireless systems will provide significantly higher data rates, offer a variety of services and applications and allow global roaming among a diverse range of mobile access networks [1].

In a typical 4G networking scenario, mobile terminals equipped with multiple interfaces have to determine the best network among the available networks. For a satisfactory user experience, mobile terminals must be able to seamlessly transfer to the best network without any interruption to an ongoing service. Such ability to handover between heterogeneous networks is referred to as Seamless Vertical Handoff (VHO) [2]. As a result, an interesting problem surfaced on how to decide the best network to use at the best time moment.

Vertical handoff provides a mobile user great flexibility for network access. However, the decision on which network to use becomes much more complicated, because both the number of networks and the decision criteria increase. Thus an intelligent vertical handover decision

(VHD) algorithm is essential for 4G network access. As the mobile users move in an environment with different networks supporting different technologies, the VHD depends on different criteria such as bandwidth, cost, power consumption, user preferences and security [3].

All the existing approaches mainly focused on the vertical handoff decision, assuming that the handoff decision processing task is performed on the mobile side. Such process requires a non negligible amount of resources to exchange information between MT and neighbor networks in order to accomplish the discovery of the best network to handoff. The main issues of the handoff decision : combining decision criteria, comparing them and answering the user needs anytime and anywhere. Several proposals and approaches considering VHD algorithms were proposed in the literature.

This paper proposes a vertical handoff decision algorithm in order to determine the best network based on dynamic factors such as RSS, Velocity and Position of the mobile terminal and static factors of each network. Thus, this algorithm meets the individual needs and also improve the whole system performance by reducing the unnecessary handoffs and increasing the throughput.

II. RELATED WORK

An efficient vertical handoff (VHO) is very essential in ensuring the system performance because the delay experienced by each handoff has a greater impact on the quality of multimedia services. The VHD algorithm should reduce the number of unnecessary handoffs to provide better throughput to all flows. Research on design and implementation of optimized VHD algorithms has been carried out by many scholars using various techniques. Based on the handoff decision criteria, VHD algorithms are categorized as RSS based algorithms, Bandwidth based algorithms, User Mobility based algorithms and Cost function based algorithms

In RSS based algorithms, RSS is used as the main criterion for handoff decision. Various schemes have been developed to compare RSS of the current point of attachment with that of the candidate point of attachments. They are: Relative RSS, RSS with hysteresis, RSS with hysteresis plus dwelling timer method [4,5]. Relative RSS is not applicable for VHD, since the RSS from different types of networks can not be compared directly due to the disparity of the technologies involved. In RSS with hysteresis method, handoff is performed whenever the RSS

of new Base station (BS) is higher than the RSS of old BS by a predefined value. In RSS with hysteresis plus dwelling timer method, whenever the RSS of new BS is higher than the RSS of old BS by a predefined hysteresis, a timer is set. When it reaches a certain specified value, handoff is processed. This minimizes Ping pong handoffs. But other criteria have not been considered in this method. EVHD algorithm makes use of this method for RSS comparison.

In bandwidth based algorithms, available Bandwidth for a mobile terminal is the main criterion. In [6], a bandwidth based VHD method is presented between WLANs and a WCDMA network using Signal to Interference and Noise ratio (SINR). It provides users higher throughput than RSS based handoffs since the available bandwidth is directly dependent on the SINR. But it may introduce excessive handoffs with the variation of the SINR. This excessive handoffs is reduced by a VHD heuristic based on the wrong decision probability (WDP) prediction [7]. The WDP is calculated by combining the probability of unnecessary and missing handoffs. This algorithm is able to reduce the WDP and balance the traffic load. But in the above papers, RSS has not been considered. A handoff to a target network with high bandwidth but weak received signal is not desirable as it may result in connection breakdown.

In user mobility based algorithms, velocity information is a critical one for handoff decision. In the overlay systems, to increase the system capacity, micro/pico cells are assigned for slow moving users and macro cells are assigned for fast moving users by using velocity information [8]. It decreases the number of dropped calls. An improved handoff algorithm [9] has been presented to reduce the number of unnecessary handoffs by using location and velocity information estimated from GSM measurement data of different signal strengths at MT received from base stations. From these papers, it is seen that velocity and location information are also having great effect on handoff management. They should also be taken into account in order to provide seamless handoff between heterogeneous wireless networks.

Cost function based algorithms combine network metrics such as monetary cost, security, power consumption and bandwidth. The handoff decision is made by comparing the result of this function for the candidate networks [10,11,12]. Different weights are assigned to different input metrics depending on the network conditions and user preferences. These algorithms have not considered other dynamic factors, such as velocity, position of the MT.

III. PROPOSED VERTICAL HANDOFF DECISION ALGORITHM

EVHD algorithm is a combined algorithm that combines the static parameters of the network such as usage cost, bandwidth and power consumption and dynamic parameters such as RSS, velocity and position of the MT. The main objective of EVHD is to maximize the throughput by reducing the number of handoffs. The EVHD algorithm involves two phases: the calculation of Gain function and the calculation of Overall Gain function.

Calculation of Gain function provides cost differentiation. The Gain function calculates the cost of the possible target network. It is a function of the offered bandwidth B, Power consumption P and usage charge of the network C.

$$G_n = f(B_n, P_n, C_n)$$

G_n is the Gain function for network n. The Gain function is calculated by using Simple Additive Weight (SAW) algorithm.

$$\text{Gain function } G_i = w_b f_{b,i} + w_p f_{p,i} + w_c f_{c,i}$$

Where w_b is weight factor for offered bandwidth, w_p is weight factor for power consumption by network interface and w_c is weight factor for the usage cost of network.

$f_{b,i}$, $f_{p,i}$ and $f_{c,i}$ represent the normalized values of network i for bandwidth, power consumption and usage cost respectively. Based on the service requirement, the weights are assigned to the parameters.

Calculation of Overall Gain function provides the best network to handoff. A candidate network is the network whose received signal strength is higher than its threshold and its position is less than the threshold. The RSS of MT is measured using the path loss and shadowing formula that is widely adopted for ns-2. The RSS of MT can be expressed as

$$RSS = PL(d_0) - 10\log(d/d_0) + X_\sigma$$

Where $PL(d_0)$ is the received power at a reference distance (d_0). The simple free space model is used to compute $PL(d_0)$. d is the distance between servicing BS and MT. n is the path loss exponent. X_σ is a Gaussian random variable with zero mean and standard deviation of σ .

Fluctuations in RSS are caused by shadowing effect. They lead the MT into unnecessary ping-pong handoffs. To avoid these ping-pong handoffs, a dwell timer is added. The timer is started when the RSS is less than RSS threshold. The MT performs a handoff if the condition is satisfied for the entire timer interval.

The position of the MT is measured. It is based on the concept that a handoff should be performed before the MT reaches a certain distance from the BS. This is known as position threshold [8].

$$\text{Position threshold } r = a - v\tau$$

Where a is radius of the service area of the BS, v is velocity of the MT and τ is estimated handoff signaling delay.

The priority for each network is based on the difference which is measured for each network.

$$RSS \text{ difference} = RSS - RSS \text{ threshold}$$

$$\text{Position diff} = \text{position threshold} - \text{position of the MT}$$

Higher the difference means higher the priority. It is so because higher difference indicates that the MT is more nearer to the BS of that network. Hence the MT can stay for more time in the cell of the respective network before asking for another handoff. Thus it is possible to reduce the unnecessary handoffs and improve the performance of the system.

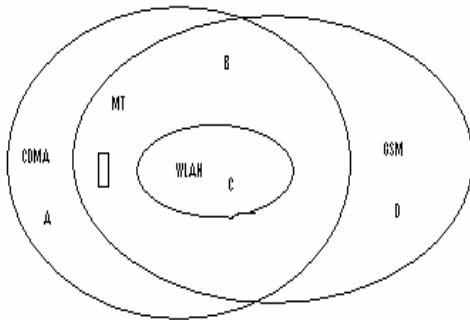
The priority levels p_i are assigned to the networks according to the difference. Overall Gain (OG) is calculated by multiplying Gain function by this priority level.

$$OG = G * p_i$$

A candidate network which has the highest overall Gain is selected as the best network to handoff.

IV. SIMULATION

A simulation model where two cellular systems GSM and CDMA and a WLAN forming an overlay structure is considered for simulation given below.



The MT can be in any one of the regions A, B, C, D. For this simulation, the following values are assigned for the parameters: $PL(d_0) = -30\text{dB}$, $n = 4$, $d_0 = 100\text{m}$, $\sigma = 8\text{dB}$, $\tau = 1\text{sec}$

	WLAN	GSM	CDMA
Offered bandwidth	2Mbps	100kbps	150kbps
Power consumption	3hrs	2.5hrs	2hrs
Usage cost	10Rs/min	5Rs/min	2.5Rs/min
RSS threshold	-60dB	-80dB	-70Db
Velocity threshold	11m/s	13m/s	12m/s

The simulation has been performed for static parameter based, RSS based, RSS-timer based and EVHD algorithms. In Static factors based algorithm, static parameters alone have been considered and hence causes lot of false handoffs. In RSS based algorithm, RSS of the MT has been compared with the signal strength threshold of the respective network. If it is lesser than the threshold, handoff is performed. But because of some shadowing effects, the signal strength is used to fluctuate and cause a lot of false handoff trigger. In RSS-timer based algorithm, RSS has been recorded over a period of time. This timer is applied to reduce the fluctuations of RSS caused by shadowing effect and hence, to reduce ping-pong handoff. In the Proposed EVHD algorithm, static parameters, RSS, velocity and position of the MT have been considered for handoff decision. A handoff is carried out whenever the position of the MT reaches to a certain boundary, regardless of RSS. This reduces the handoff failure. The boundary is a safety distance of MT from the BS to assure a successful handoff and this boundary is not fixed and is varying according to the position and velocity of MT.

V. RESULTS AND DISCUSSION

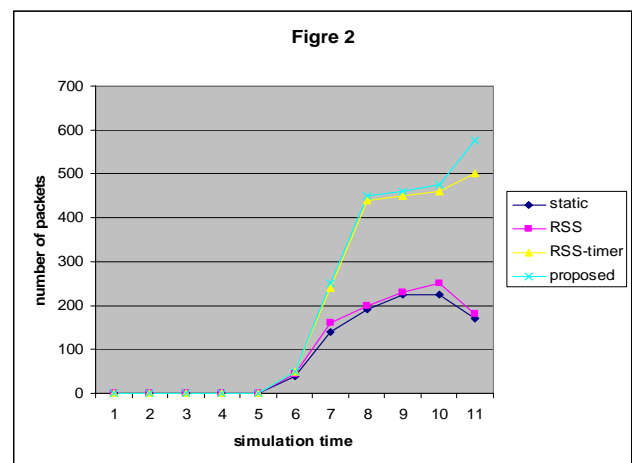
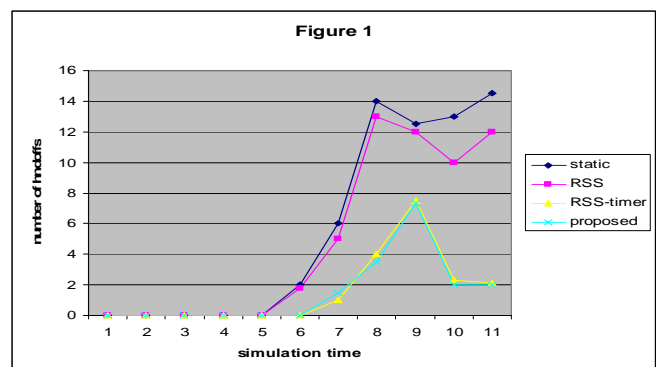
In this study, the performance evaluation of the efficient vertical handoff decision algorithm (EVHD) has been carried out and the metrics number of unnecessary handoffs and throughput have been compared with static parameter based algorithm, RSS-static parameter based algorithm, RSS-timer-static parameter based algorithm.

The number of handoffs experienced by the algorithms is shown in fig.1. The obtained results show that the number

of handoffs is reduced by 78.5% in EVHD algorithm compared to static parameter based algorithm and 25% compared to RSS-timer based algorithm. The huge reduction in number of handoffs, is one of the major achievements of EVHD algorithm.

The number of packets serviced by static parameter based, RSS based, RSS-timer based and EVHD algorithms have been observed and is shown in fig.2. EVHD algorithm is able to service more number of packets in a given period of time compared to other algorithms because of its reduction in number of handoffs.

The results show that the EVHD algorithm performs better in terms of number of handoffs and throughput compared to static parameter based, RSS based, RSS-timer based algorithms.



VI. CONCLUSION

Efficient vertical handoff decision algorithm is a combined algorithm that combines the static parameters of the network such as usage cost, bandwidth and power consumption and dynamic parameters such as RSS, velocity and position of the MT. The algorithm has been implemented successfully using ns-2 simulator. The results show that EVHD achieves about 78.5% reduction in number of handoffs compared to static parameter based algorithm and 25% reduction compared to RSS-timer based algorithm and it is clear that EVHD provides better throughput with minimum number of handoffs compared to other algorithms. Thus EVHD has outperformed the other algorithms by providing less number of handoffs and high throughput and hence it is efficient in enhancing QoS for multimedia applications.

REFERENCES

- [1] M.Kassar, B.Kervelia, G.Pujolle, An overview of vertical handover decision strategies in heterogeneous wireless networks, *Computer communications* 31 (10) (2008)
- [2] J. McNair, F. Zhu, Vertical handoffs in fourth-generation multinetwork environments, *IEEE Wireless Communications* 11 (3) (2004) 8–15.
- [3] N. Nasser, A. Hasswa, H. Hassanein, Handoffs in fourth generation heterogeneous networks, *IEEE Communications Magazine* 44 (10) (2006) 96–103.
- [4] S. Mohanty, I.F. Akyildiz, A cross-layer (layer 2 + 3) handoff management protocol for next-generation wireless systems, *IEEE Transactions on Mobile Computing* 5 (10) (2006) 1347–1360.
- [5] A.H. Zahran, B. Liang, A. Saleh, Signal threshold adaptation for vertical handoff in heterogeneous wireless networks, *Mobile Networks and Applications* 11 (4) (2006) 625–640.
- [6] K. Yang, I. Gondal, B. Qiu, L.S. Dooley, Combined SINR based vertical handoff algorithm for next generation heterogeneous wireless networks, in: *Proceedings of the 2007 IEEE Global Telecommunications Conference (GLOBECOM'07)*, Washington, DC, USA, November 2007, pp. 4483–4487.
- [7] C. Chi, X. Cai, R. Hao, F. Liu. Modeling and analysis of handover algorithms, in: *Proceedings of the 2007 IEEE Global Telecommunications Conference (GLOBECOM'07)*, Washington, DC, USA, November 2007, pp. 4473–4477.
- [8] Xiao.C, K.D.Mann and J.C.Olivier, 2001, Mobile speed estimation for TDMA-based hierarchical cellular systems, *Proc.Trans. Veh.Technol*, 50, 981-991.
- [9] Juang.R.T, H.P.Lin and D.B.Lin, 2005, An improved location-based handover algorithm for GSM systems, *Proc of wireless communications and networking conference*, Mar.13-17, pp-1371-1376.
- [10] A. Hasswa, N. Nasser, H. Hassanein, Tramcar: a context-aware crosslayer architecture for next generation heterogeneous wireless networks, in: *Proceedings of the 2006 IEEE International Conference on Communications (ICC'06)*, Istanbul, Turkey, June 2006, pp. 240–245.
- [11] R. Tawil, G. Pujolle, O. Salazar, A vertical handoff decision scheme in heterogeneous wireless systems, in: *Proceedings of the 67th Vehicular Technology Conference (VTC'08 – Spring)*, Marina Bay, Singapore, April 2008, pp. 2626–2630.
- [12] F. Zhu, J. McNair, Optimizations for vertical handoff decision algorithms, in: *Proceedings of the 2004 IEEE Wireless Communications and Networking Conference (WCNC'04)*, Atlanta, Georgia, USA, March 2004, pp. 867–872.

AUTHORS PROFILE

Mrs. S.Aghalya has received her B.E. degree in Electronics and Communication Engineering from Madras university , India in 1991 and M.E. degree in Optical Communication from Anna University ,India in 2001. She is an Assistant Professor at St.Joseph's College Of Engg, Chennai ,India. She has 16 years of Teaching experience. She is currently pursuing her research at Anna University Trichy, India . Her research interest is in wireless networks.

P. Seethalakshmi has received her B.E. degree in Electronics and Communication Engineering in 1991 and M.E. degree in Applied Electronics in 1995 from Bharathiar University, India. She obtained her doctoral degree from Anna University Chennai, India in the year 2004. She has 15 years of Teaching experience and her areas of research includes Multimedia Streaming, Wireless Networks, Network Processors and Web Services.

COMBINING LEVEL- 1 ,2 & 3 CLASSIFIERS FOR FINGERPRINT RECOGNITION SYSTEM

Dr.R.Seshadri ,B.Tech,M.E,Ph.D
Director, S.V.U.Computer Center
S.V.University, Tirupati

Yaswanth Kumar.Avulapati,M.C.A,M.Tech,(Ph.D)
Research Scholar, Dept of Computer Science
S.V.University, Tirupati

Abstract

Biometrics is the science of establishing the identity of an person based on their physical, chemical and behavioral characteristics of the person. Fingerprints are the most widely used biometric feature for person identification and verification in the field of biometric identification .A finger print is the representation of the epidermis of a finger. It consists of a pattern of interleaved ridges and valleys.

Fingerprints are graphical flow-like ridges present on human fingers. They are fully formed at about seven months of fetus development and finger ridge configurations do not change throughout the life of an individual except due to accidents such as bruises and cuts on the fingertips.

This property makes fingerprints a very attractive biometric identifier. Now a day's fingerprints are widely used among different biometrics technologies. In this paper we proposed an approach to classifying the fingerprints into different groups. These fingerprints classifiers are combined together for recognizing the people in an effective way.

Keywords-Biometrics, Classifier,Level-1,Level-2 features,Level-3 features

Introduction

A fingerprint is a pattern of ridges and valleys located on the tip of each finger. Fingerprints were used for personal identification for many centuries and the matching accuracy was very high. Human fingerprint recognition has a tremendous potential in a wide variety of forensic, commercial and law enforcement applications.

Fingerprints are broadly classified into three levels they are Level-1 which includes arch,tentarch, loop, double Loop, pocked Loop, whorl ,mixed, left-loop, right-loop the Level-2 includes the minutiae and Level 3 includes pores etc.

There are so many approaches are there for recognizing the fingerprints among these correlation based, minutiae based, ridge feature based are most popular ones.

Several biometrics systems have been successfully developed and installed. However some methods do not perform well in many real-world situations due to its noise.

Fingerprint Classifier

Here we proposed a fingerprint classifier framework. A combination scheme involving different fingerprint classifiers which integrates vital information is likely to improve the overall system performance.

The fingerprint classifier combination can be implemented at two levels feature level and decision level. We use the decision level combination that is more appropriate when the component classifiers use different types of features. Kittler provides a theoretical framework to combine various classifiers at the decision level. Many practical applications of combining multiple classifiers have been developed. Brunelli and Falavigna presented a person identification system by combining outputs from classifiers based on Audio and visual.

Here the combination approach is designed at the decision level utilizing all the available information, i.e. a subset of (Fingerprint) labels along with a confidence value, called the matching score provided by each of the nine finger print recognition method.

Classification of Fingerprint (Level-1, Level -2 & Level-3) Features

Level 1 features describe the ridge flow pattern of a fingerprint. According to the Henry classification system there are eight major pattern classes, comprised of whorl, left loop, right loop, twin loop, arch, tented arch. as shown in the figure 1.

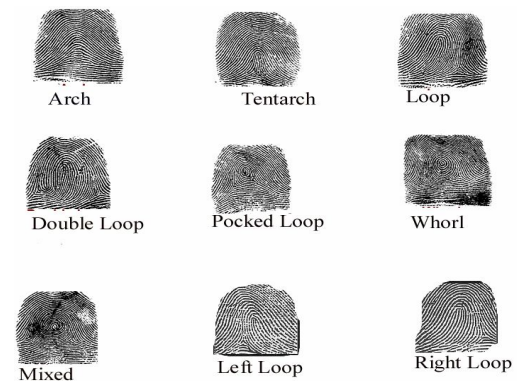


Fig 1. Fingerprint Level 1 Features

Level 2 features describe various ridge path deviations where single or multiple ridges form abrupt stops, splits, spurs bifurcation Composite minutiae (i.e., forks, spurs, bridges, crossovers and bifurcations) can all be considered as combinations of these basic forms enclosures, etc. These features, known as the Galton points or minutiae, have two basic forms: ridge ending and ridge as shown in fig 2.

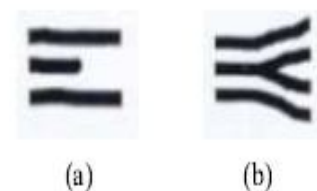


Fig 2. Fingerprint Level 2 Features

Level 3 features refer to all dimensional attributes of a ridge, such as ridge path deviation, width, shape, pores, edge contour, incipient ridges, breaks, creases, scars and other permanent details as shown in fig 3.

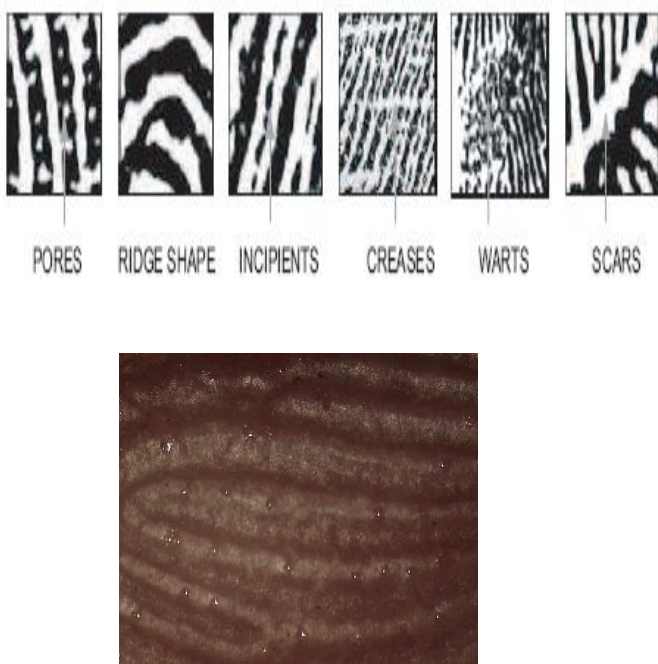


Fig 3. Fingerprint Level 3 Features

Classifier Combination System

We proposed a classifier combination system shown in the Fig. Here currently we use only nine classifiers for level-1 features of fingerprints namely arch, tent arch, loop, double Loop, pocked Loop, whorl, mixed, left-loop, right-loop

For Fingerprint level-2 features namely right-loop various ridge path deviations where single or multiple ridges form abrupt stops, splits, spurs bifurcation Composite minutiae (i.e., forks, spurs, bridges, crossovers and bifurcations

For Level-3 features namely deviation, width, shape, pores, edge contour, incipient ridges, breaks, creases, scars

Following two strategies are provided for integrating outputs of individual classifiers, (i) the sum rule, and (ii) a RBF network as a classifier, using matching scores¹³⁰ as the input feature vectors as shown in fig 4.

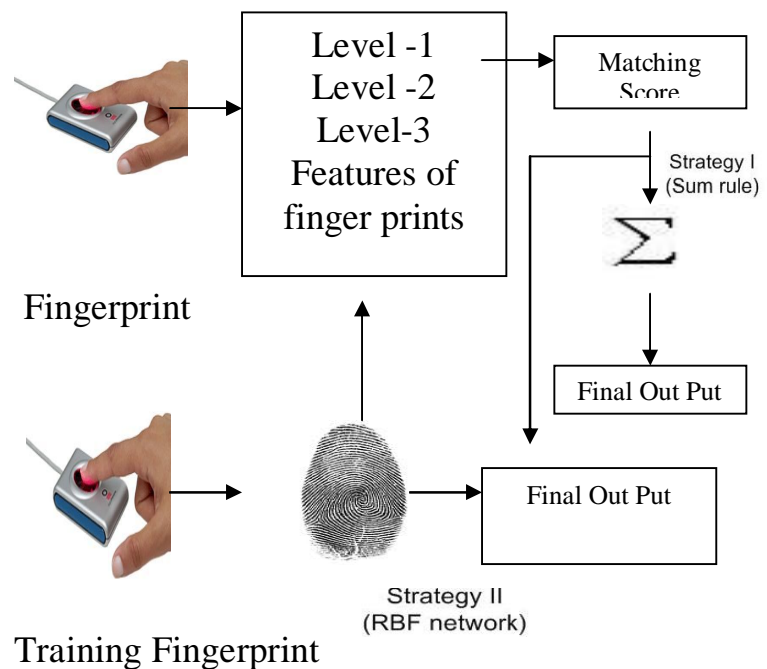


Fig 4. Fingerprint Classifier Combination System

Combination Strategy

Kittler analyzed several classifier combination rules and concluded that the sum rule as shown in the given below outperforms other combination schemes based on empirical observations.

Unlike explicitly setting up combination rules, it is possible to design a new classifier using the outputs of individual classifiers as features to this new classifier.

Here we assume the RBF network as a new classifier. Given m templates in the training set, m matching scores will be output for each test image from each classifier. We consider the following two integration strategies

1. Strategy I: Sum Rule. The combined matching score is calculated as

$$\text{Macomb} = \text{MSPCA} + \text{MSICA} + \text{MSLDA}:$$

For a given sample, Output the class with the largest value of *Macomb*.

2. Strategy II: RBF network. For each test image, the *m* matching scores obtained from each classifier are used as a feature vector. Concatenating these feature vectors derived from Level-1, Level-2, Level-3 classifiers results in a feature vector of size *3m*.

An RBF network is designed to use this new feature vector as the input to generate classification results. We adopt a Level-1, Level-2, Level-3 layers RBF network. The input layer has 3 levels *m* nodes and the output has *c* nodes, where *c* is the total number of classes (number of distinct features of fingerprints). In the output layer, the class corresponding to the node with the maximum output is assigned to the input image. The number of nodes in the hidden layer is constructed empirically, depending on the sizes of the input and output layers. Sum score is output as the final result.

The recognition accuracies of different finger print recognition approaches are listed in table 5a. The cumulative match score vs. rank curve is used to show the performance of each classifier, see Fig 5b. Since our RBF network outputs the final label, no rank information is available. As a result, we cannot compute the cumulative match score vs. rank curve for RBF combination

Level-1 Features	Level-2 Features	Level-3 Features
70	75	90

Fig.5a recognition accuracies of different finger print recognition approaches are listed

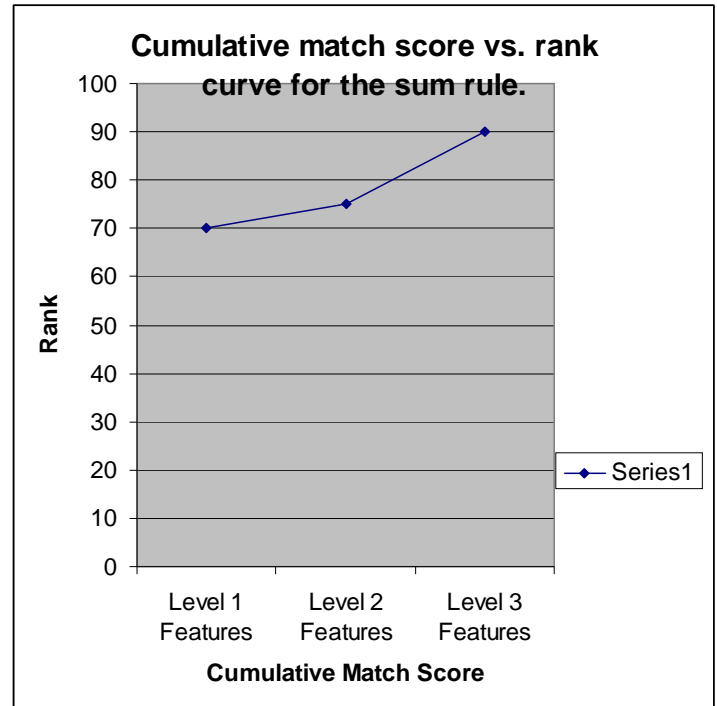


Figure 5 b show that the combined classifiers, based on both the sum-rule and RBF network, outperform each individual classifier.

CONCLUSION

Finally we conclude that in our proposed approach the combination scheme which combines the output matching scores of three levels of well-known Fingerprint recognition system. Basically we proposed the model to improve the performance of a fingerprint identification system at the same time the system provides high security from unauthorized access.

Two mixing strategies, sum rule and RBF-based integration are implemented to combine the output information of three level features of fingerprint individual classifiers.

The proposed system framework is scalable other fingerprint recognition modules can be easily added into this framework. Results are encouraging, illustrating that both the combination strategies lead to more accurate fingerprint recognition than that made by any one of the individual classifiers.

References

- [1].A. K. Jain,Patrick Flynn,Arun A.Ross .
“Handbook of Biometrics”.
- [2].D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, Handbook of Fingerprint Recognition. Springer, 2003.
- [3.] N. Yager and A. Amin. Fingerprint classification: A review. Pattern Analysis Application, 7:77{93, 2004.
- [4]. O. Yang, W. Tobler, J. Snyder, and Q. H. Yang. Map Projection Transformation. Taylor & Francis, 2000.
- [5]. Z. Zhang. Flexible Camera Calibration by Viewing A Plane from Unknown Orientations. IEEE Transactions on Pattern Analysis and Machine Intelligence,11:1330{1334, 2000.
- [6]. J. Zhou, C. Wu, and D. Zhang. Improving Fingerprint Recognition Based on Crease Detection. In Proc. International Conference on Biometric Authentication (ICBA), pages 287{293, Hong Kong, China, July 2004.
- [7]. Y. Zhu, S. Dass, and A. K. Jain. Statistical Models for Assessing the Individuality of Fingerprints. IEEE Transactions on Information Forensics and Security, 2:391{401, 2007.
- [8]. Y. F. Zhu. Statistical Models for¹³² Fingerprint Individuality. PhD thesis, Depart-

ment of Computer Science and Engineering, Michigan State University, 2008.

- [9].K. Kryszczuk, A. Drygajlo, and P. Morier. Extraction of Level 2 and Level 3 Features for Fragmentary Fingerprints. In Proc. COST Action 275 Workshop, pages 83{88, Vigo, Spain, 2004.

- [10]A. K. Jain, S. Prabhakar, and S. Chen, “Combining multiple matchers for a high security fingerprint verification system,” *Pattern Recognition Letters*, vol. 20, no. 11-13, pp. 1371–1379, 1999.

Authors Profile



Dr.R.Seshadri was born in Andhra Pradesh, India, in 1959. He received his **B.Tech** degree from Nagarjuna University in 1981. He received his **M.E** degree in Control System Engineering from PSG College of Technology, Coimbatore in 1984. He was awarded with **PhD** from Sri Venkateswara University, Tirupati in 1998. He is currently Director, Computer Center, S.V.University, Tirupati, India. He has Published number of papers in national and international conferences, seminars and journals. At present 12 members are doing research work under his guidance in different areas



Mr.YaswanthKumar .Avulapati received his **MCA** degree with **First class** from Sri Venkateswara University, Tirupati. He received his **M.Tech** Computer Science and Engineering degree with **Distinction** from Acharya Nagarjuna University, Guntur.He is a research scholar in S.V.University Tirupati, Andhra Pradesh.He has presented number of papers in national and international conferences, seminars.He attend Number of work shops in different fields.

Preventing Attacks on Fingerprint Identification System by Using Level-3 Features

Dr.R.Seshadri ,B.Tech,M.E,Ph.D
Director, S.V.U.Computer Center
S.V.University, Tirupati

Yaswanth Kumar.Avulapati,M.C.A,M.Tech,(Ph.D)
Research Scholar, Dept of Computer Science
S.V.University, Tirupati

Abstract

Biometrics is the science of establishing the identity of an individual based on their physical, behavioral and chemical characteristics of the person. Fingerprints are the most widely used biometric feature for person identification and verification in the field of biometric identification .

A finger print is the representation of the epidermis of a finger. It consists of a pattern of interleaved ridges and valleys.

Now a days Fingerprints are widely used technique among other biometric like Iris,Gait,Hand Geometry, Dental Radiographs etc. Fingerprint Ridges, Minutae and sweat pores do not change throughout the life of an human being except due to accidents such as bruises and cuts on the fingertips.

This property makes fingerprints a very attractive biometric identifier. In this paper we proposed a biometrics system which

Prevents from Attacks from Gummy fingerprints. We proposed Fingerprint Identification System which is immune to attacks by Using Level-3 Features

Keywords-

Biometrics, Immune, Sweat pores, Level-3 features

Introduction

A fingerprint is a pattern of ridges and valleys located on the tip of each finger. Fingerprints were used for personal identification for many centuries and the matching accuracy was very high.Now a days the possible threats caused by something like real fingers, which are called fake or artificial fingers, should be critical for authentication based on fingerprint system

Conventional fingerprint systems cannot categorize between an impostor who falsely obtains the access privileges from a ATM system or any other source (e.g., secrete key,

passwords) of a genuine user and the genuine user. Moreover biometric systems (Ex. Fingerprint identification system) can be more suitable for the users since there is no secret keys, password to be forgotten and a fingerprint identification system (biometric system) can be used to access several applications without any trouble of remembering passwords.

There are many advantages by using the biometric systems. These systems are in danger to attacks which can decrease their security. According to Ratha et al. analyzed these attacks, and grouped them into eight classes. These attacks along with the components of a typical biometric system that can be compromised. Type 1 attack involves presenting a fake biometric (e.g., synthetic fingerprint, face, iris) to the sensor. Submitting a previously intercepted biometric data constitutes

The second type of attack (replay). In the third type of attack, the feature extractor module is compromised to produce feature values selected by the attacker. Genuine feature values are replaced with the ones selected by the attacker in the fourth type of attack. Matcher can be modified to output an artificially high matching score in the fifth type of attack. The attack on the template database (e.g., adding a new template, modifying an existing template, removing templates, etc.) constitutes the sixth type of attack.

The transmission medium between the template database and matcher is attacked in the seventh type of attack, resulting in the alteration of the transmitted templates. Finally, the matcher result (accept or reject) can be overridden by the attacker.

According to Matsumoto et al. he attacked 11 different fingerprint verification systems with artificially created gummy (gelatin) fingers. For a cooperative owner, her finger is pressed to a plastic mold, and gelatin leaf is used to create the gummy finger. The operation is said to take less than an hour. It was found that the gummy fingers could be enrolled in all of the 11 systems, and they were accepted with a probability of 68-100%. When the owner does not cooperate, a residual fingerprint from a glass plate is enhanced with a cyanoacrylate adhesive.

After capturing an image of the print, PCB based processing similar to the operation described above is used to create the gummy fingers. All of the 11 systems enrolled the gummy fingers and they accepted the gummy fingers with more than 67% probability.

Threat investigation for Fingerprint Identification system Systems

Fingerprint identification systems capture the fingerprints and extract fingerprint features from the images, encrypt the features, transmit them on communication media and then store them as templates into database. Some systems encrypt templates with a secure cryptographic scheme and manage not whole original images but compressed images. Therefore, it is said to be difficult to reproduce valid fingerprints by using the templates. Some systems are secured against a so-called replay attack in which an attacker copies a data stream from a fingerprint scanner to a server and later replays it with an one time protocol or a random challenge response device.

When a valid user has registered his/her live finger with a fingerprint identification system there would be several ways to mislead the system. In order to mislead the fingerprint

identification system an attacker may present the following things to its fingerprint scanner

- 1) The registered finger
- 2) An unregistered finger (Imposter's finger)
- 3) A genetic clone of the registered finger
- 5) An artificial clone of the registered finger

Making a Artificial Fingerprint from a live fingerprint

Materials needed for making Gummy fingerprint as shown in fig1.

a) Free Molding plastic “Free plastic”

b) Solid Gelatine sheet from



200JPY/30grams



Free plastic and Gelatin sheet

Fig .1.Materials used for Fake Fingerprints

How they are making a mold

Here we present how an attacker making a gummy fingerprint as shown in the fig.2.The following steps for making the gummy fingerprint. It takes up to 10 min.

- 1) Put the plastic in hot water to soften it
- 2) Press a live finger against it
- 3) The mold



Put the plastic in hot water to soften it



Press a live finger against it



The mold

Fig.2. Steps for Making the Fake Fingerprints

Preparation of material

The preparation of gelatin liquid steps is shown in fig.3.

Step1:-The liquid in which immersed gelatin at 50 wt

Step 2:- Adding the hot water 30cc to solid gelatin 30 grm



Pour Gelatin into bottle



Pour Hot water into bottle



Steer the bottle with gelatin

Fig.3. Preparation Of Gelatin Liquid

Preparation Of Fake (Or) Gummy Fingerprint



Pour the liquid into mold



Put the mold into refrigerator for cool



The Fake or Gummy Fingerprint for use

Fig.4.Preparation of Fake or Gummy Fingerprint

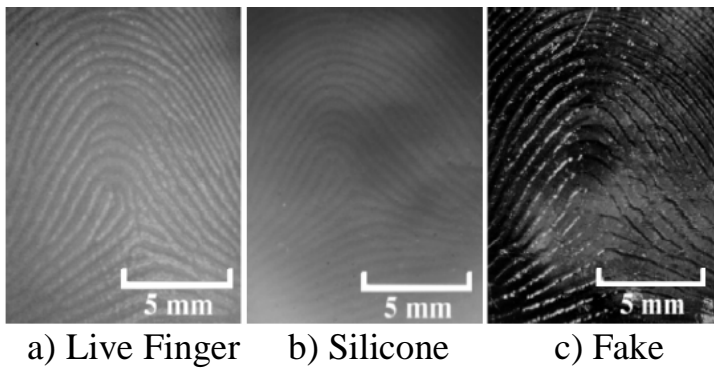


Fig.5.Comparison of Live and Fake Fingerprint (Similarities is there)

Proposed system:

The proposed a biometrics system which prevents from Attacks from Fake (or)Gummy fingerprints. We proposed Fingerprint Identification System which is immune to attacks by Using Level-3 Features as shown in the fig 6.

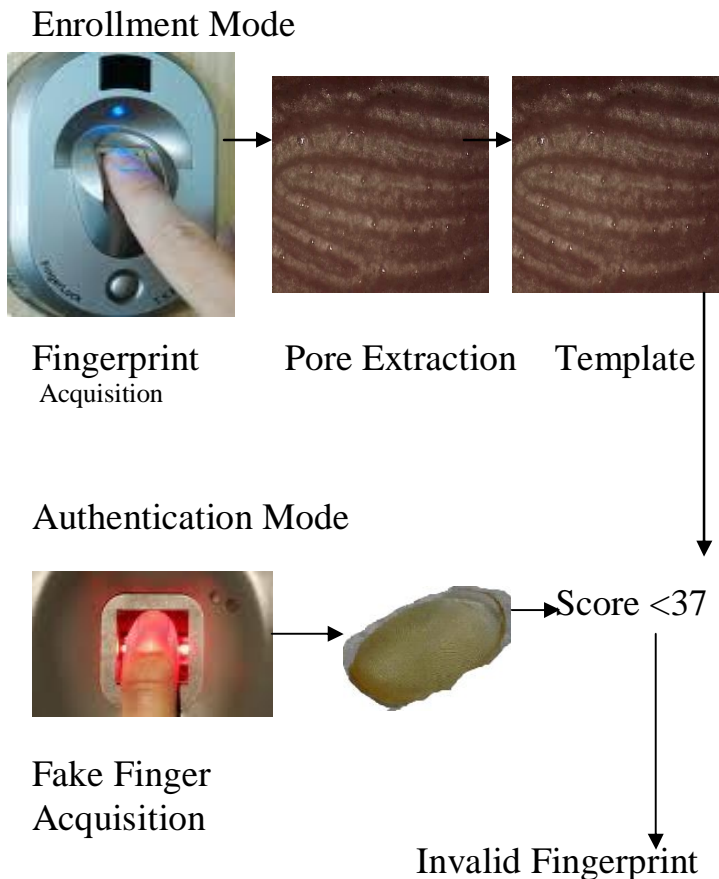


Fig.6.Enrollment of Guanine and Authentication mode (Fake) fingerprint system

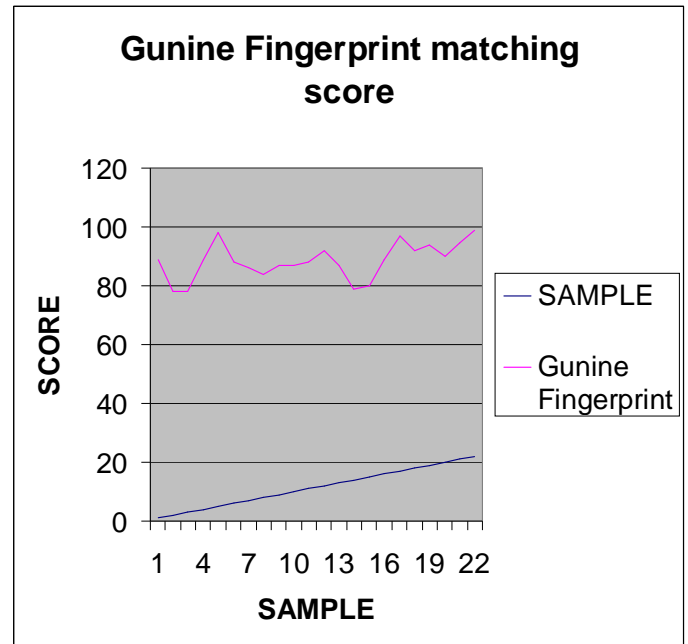


Fig.7. Shows the Guanine Fingerprint Matching Score using Level-3 Features

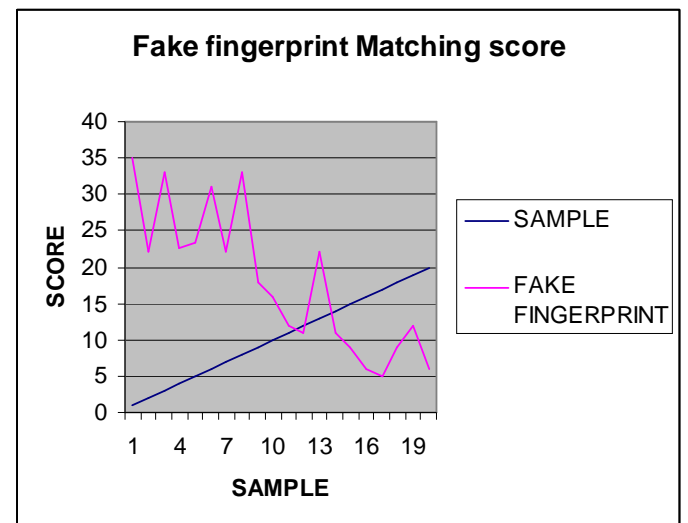


Fig.8. Shows the Fake Fingerprint Matching Score using Level-3 Features

Coclusion:

This paper presents an approach to immune biometric systems which Prevents from Attacks from Fake (or) Gummy fingerprints. There can be various attacks using the fake (or) Gummy finger prints. Now a days fake fingerprints are easy to

prepare easily obtained materials with low cost. The manufacturers and the users of biometrics systems should carefully examine the security of the system against fake fingerprints. Here we proposed Fingerprint Identification System which is immune to attacks by Using Level-3 Features like Pores etc.

References

- [1]. Extended Feature Set and Touchless Imaging For Fingerprint Matching By Yi Chen A Dissertation Su 2009
- [2]. N. Yager and A. Amin. Fingerprint classification: A review. Pattern Analysis Application, 7:77{93, 2004.
- [3]. O. Yang, W. Tobler, J. Snyder, and Q. H. Yang. Map Projection Transformation. Taylor & Francis, 2000.
- [4]. Z. Zhang. Flexible Camera Calibration by Viewing A Plane from Unknown Orientations. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11:1330{1334, 2000.
- [5]. J. Zhou, C. Wu, and D. Zhang. Improving Fingerprint Recognition Based on Crease Detection. In Proc. International Conference on Biometric Authentication (ICBA), pages 287{293, Hong Kong, China, July 2004.
- [6]. Y. Zhu, S. Dass, and A. K. Jain. Statistical Models for Assessing the Individuality of Fingerprints. IEEE Transactions on Information Forensics and Security, 2:391{401, 2007.
- [7]. Y. F. Zhu. Statistical Models for Fingerprint Individuality. PhD thesis, Department of Computer Science and Engineering, Michigan State University, 2008.
- [7]. SWGFAST. Scientific Working Group on Friction Ridge Analysis, Study and Technology. <http://www.swgfast.org/>, 2006.

- [8]. A. K. Jain, A. Nagar, and K. Nandakumar. Latent Fingerprint Matching. Technical Report MSU-CSE-07-203, Michigan State University, 2007.

Authors Profile



Dr. R. Seshadri was born in Andhra Pradesh, India, in 1959. He received his **B.Tech** degree from Nagarjuna University in 1981. He received his **M.E** degree in Control System Engineering from PSG College of Technology, Coimbatore in 1984. He was awarded with **PhD** from Sri Venkateswara University, Tirupati in 1998. He is currently Director, Computer Center, S.V. University, Tirupati, India. He has Published number of papers in national and international conferences, seminars and journals. At present 12 members are doing research work under his guidance in different areas



Mr. Yaswanth Kumar .Avulapati received his **MCA** degree with **First class** from Sri Venkateswara University, Tirupati. He received his **M.Tech** Computer Science and Engineering degree with **Distinction** from Acharya Nagarjuna University, Guntur. He is a research scholar in S.V. University Tirupati, Andhra Pradesh. He has presented number of papers in national and international conferences, seminars. He attend Number of work shops in different fields.

Using Fuzzy Support Vector Machine in Text Categorization Base on Reduced Matrices

Vu Thanh Nguyen¹

¹University of Information Technology HoChiMinh City, VietNam

Abstract - In this article, the authors present result compare from using Fuzzy Support Vector Machine (FSVM) and Fuzzy Support Vector Machine which combined Latin Semantic Indexing and Random Indexing on reduced matrices (FSVM_LSI_RI). Our results show that FSVM_LSI_RI provide better results on Precision and Recall than FSVM. In this experiment a corpus comprising 3299 documents and from the Reuters-21578 corpus was used.

Keyword – SVM, FSVM, LSI, RI

I. INTRODUCTION

Text categorization is the task of assigning a text to one or more of a set of predefined categories. As with most other natural language processing applications, representational factors are decisive for the performance of the categorization. The incomparably most common representational scheme in text categorization is the Bag-of-Words (BoW) approach, in which a text is represented as a vector t of word weights, such that $t_i = (w_1 \dots w_n)$ where w_n are the weights of the words in the text. The BoW representation ignores all semantic or conceptual information; it simply looks at the surface word forms. BoW modern is based on three models: Boolean model, Vector Space model, Probability model.

There have been attempts at deriving more sophisticated representations for text categorization, including the use of n-grams or phrases (Lewis, 1992; Dumais et al., 1998), or augmenting the standard BoW approach with synonym clusters or latent dimensions (Baker and Mc- Callum, 1998; Cai and Hofmann, 2003). However, none of the more elaborate representations manage to significantly outperform the standard BoW approach (Sebastiani, 2002). In addition to this, they are typically more expensive to compute.

In order to do this, the authors introduce a new method for producing concept-based representations for natural language data. This method is a combination of Random indexing(RI) and Latin Semantic Indexing (LSI), computation time for Singular Value Decomposition on a RI reduced matrix is almost halved compared to LSI. The authors use this method to create concept-based representations for a standard text categorization problem, and the representations as input to a FSVM

classifier. The categorization results are compared to those reached using standard BoW representations by Vector Space Model (VSM), and the authors also demonstrate how the performance of the FSVM can be improved by combining representations.

II. VECTOR SPACE MODEL (VSM) ([14]).

1. Data Structuring

In Vector space model, documents are represented as vectors in t -dimensional space, where t is the number of indexed terms in the collection. Function to evaluate terms weight:

$$w_{ij} = l_{ij} * g_i * n_j$$

- l_{ij} denotes the local weight of term i in document j .
- g_i is the global weight of term i in the document collection
- n_j is the normalization factor for document j .

$$l_{ij} = \log (1 + f_{ij})$$

$$g_i = 1 + \left(\frac{\sum_j (p_{ij} \log p_{ij})}{\log n} \right)$$

Where:

- f_{ij} is the frequency of token i in document j .

- $p_{ij} = f_{ij} / \sum_j f_{ij}$ is the probability of token i occurring in document j .

2. Term document matrix

In VSM is implemented by forming term-document matrix. Term- document matrix is $m \times n$ matrix where m is number of terms and n is number of documents.

$$A = \begin{pmatrix} d_{11} & d_{21} & \bullet & \bullet & \bullet & d_{1n} \\ d_{12} & d_{22} & \bullet & \bullet & \bullet & d_{2n} \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ d_{m1} & d_{m2} & \bullet & \bullet & \bullet & d_{mn} \end{pmatrix}$$

where:

- term: row of term-document matrix.
- document: column of term-document matrix.

- d_{ij} : is the weight associated with token i in document j .

III. LATENT SEMANTIC INDEXING (LSI) ([1]-[4])

The vector space model is presented in section 2 suffers from the *curse of dimensionality*. In other words, as the problem of sizes increase may become more complex, the processing time required to construct a vector space and query throughout the document space will increase as well. In addition, the vector space model exclusively measures term co-occurrence—that is, the inner product between two documents is nonzero if and only if there exist at least one shared term between them. Latent Semantic Indexing (LSI) is used to overcome the problems of *synonymy* and *polysemy*.

1. Singular Value Decomposition (SVD) ([5]-[9])

LSI is based on a mathematical technique called Singular Value Decomposition (SVD). The SVD is used to process decomposes a term-by-document matrix A into three matrices: a term-by-dimension matrix, U , a singular-value matrix, Σ , and a document-by-dimension matrix, V^T . The purpose of analysis the SVD is to detect semantic relationships in the documents collection. This decomposition is performed as following:

$$A = U \Sigma V^T$$

Where:

- U orthogonal $m \times m$ matrix whose columns are left singular vectors of A

- Σ diagonal matrix on whose diagonal are singular values of matrix A in descending order

- V orthogonal $n \times n$ matrix whose columns are right singular vectors of A .

To generate a rank- k approximation A_k of A where $k \ll r$, each matrix factor is truncated to its first k columns. That is, A_k is computed as:

$$A_k = U_k \Sigma_k V_k^T$$

Where:

- U_k is $m \times k$ matrix whose columns are first k left singular vectors of A

- Σ_k is $k \times k$ diagonal matrix whose diagonal is formed by k leading singular values of A

- V_k is $n \times k$ matrix whose columns are first k right singular vectors of A

In LSI, A_k is approximation of A is created and that is very important: detected a combination of literatures between terms used in the documents, excluding the change in usage term bad influence to the method to search for the index [6], [7], [8]. Because use of k -dimensional LSI ($k \ll r$) the difference is not important in the "means" is removed. Keywords often appear together in the document is nearly the same performance space in k -dimensional LSI, even the index does not appear simultaneously in the same document.

2. Drawback of the LSI model

- SVD often treated as a "magical" process.

- SVD is computationally expensive.
- Initial "huge matrix step"
- Linguistically agnostic.

IV. RANDOM INDEXING (RI) ([6],[10])

Random Indexing is an incremental vector space model that is computationally less demanding (Karlgrén and Sahlgrén, 2001). The Random Indexing model reduces dimensionality by, instead of giving each word a whole dimension, it gives them a random vector by a much lesser dimensionality than the total number of words in the text.

Random Indexing differs from the basic vector space model in that it doesn't give each word an orthogonal unit vector. Instead each word is given a vector of length 1 in a random direction. The dimension of this randomized vector will be chosen to be smaller than the amount of words in the document, with the end result that not all words will be orthogonal to each other since the rank of the matrix won't be high enough. This can be formulated as $AT = A^*$ where A is the original matrix representation of the $d \times w$ word document matrix as in the basic vector space model, T is the random vectors as a $w \times k$ matrix representing the mapping between each word w_i and the k -dimensional random vectors, A^* is A projected down into $d \times k$ dimensions. A query is then matched by first multiplying the query vector with T , and then finds the column in A^* that gave the best match. T is constructed by, for each column in T , each corresponding to a row in A , electing n different rows. $n/2$ of these are assigned the value $1/n$, and the rest are assigned $-1/n$. This ensures unit length, and that the vectors are distributed evenly in the unit sphere of dimension k (Sahlgrén, 2005).

An even distribution will ensure that every pair of vectors has a high probability to be orthogonal. Information is lost during this process (pigeonhole principle, the fact that the rank of the reduced matrix is lower). However, if used on a matrix with very few nonzero elements, the induced error will decrease as the likelihood of a conflict in each document, and between documents, will decrease. Using Random Indexing on a matrix will introduce a certain error to the results. These errors will be introduced by words that match with other words, i.e. the scalar product between the corresponding vectors will be $\neq 0$. In the matrix this will show either that false positive matches are created for every word that have a nonzero scalar product of any vector in the vector room of the matrix. False negatives can also be created by words that have corresponding vectors that cancel each other out.

Advantages of Random Indexing

- Based on Pentti Kanerva's theories on Sparse Distributed Memory.
- Uses distributed representations to accumulate context vectors.
- Incremental method that avoids the "huge matrix step".

V. COMBINING RI AND LSI

We have seen the advantages and disadvantages for both RI and LSI: RI is efficient in terms of computational time but does not preserve as much information as LSI; LSI, on the other hand, is computationally expensive, but produces highly accurate results, in addition to capturing the underlying semantics of the documents. As mentioned earlier, a hybrid algorithm was proposed that combines the two approaches to benefit from the advantages of both algorithms. The algorithm works as follows:

- First the data is pre-processed with RI to a lower dimension $k1$.
- Then LSI is applied on the reduced, lower-dimensional data, to further reduce the data to the desired dimension, $k2$.

This algorithm supposedly will improve running time for LSI, and accuracy for RI. As mentioned earlier, the time complexity of SVD D is $O(cmn)$ for large, sparse datasets. It is reasonable, then, to assume that a lower dimensionality will result in faster computation time, since it's dependent of the dimensionality m .

VI. TEXT CATEGORIZATION

1. Support Vector Machines

Support vector machine is a very specific class of algorithms, characterized by the use of kernels, the absence of local minima, the sparseness of the solution and the capacity control obtained by acting on the margin, or on other "dimension independent" quantities such as the number of support vectors.

Let $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$ a training sample set, where $x_i \in \mathbb{R}^n$ and corresponding binary class labels $y_i \in \{1, -1\}$. Let ϕ is a non-linear mapping from original data space to a high-dimensional feature space, therefore, we replace sample points x_i and x_j with their mapping images $\phi(x_i)$ and $\phi(x_j)$ respectively.

Let the weight and the bias of the separating hyper-plane is w and b , respectively. We define a hyper-plane which might act as decision surface in feature space, as following.

$$\sum_{j=1}^l w_j \phi_j(x) + b = 0$$

To separate the data linearly in the feature space, the decision function satisfies the following constrain conditions. The optimization problem is Minimize:

$$\phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i$$

Subject to:

$$y_i [(w \cdot x_i) + b] \geq 1 - \xi_i$$

$$\xi_i \geq 0, \quad i = 1, 2, \dots, l$$

where ξ_i is a slack variable introduced to relax the hard margin constraints and the regularization constant $C > 0$ implements the trade-off between the maximal margin of separation and the classification error.

To resolve the optimization problem, we introduce the following Lagrange function.

$$L = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i$$

$$- \sum_{i=1}^l \alpha_i [y_i (x_i \cdot w + b) - 1 + \xi_i]$$

$$- \sum_{i=1}^l \beta_i \xi_i$$

Where $\alpha_i \geq 0$, $\beta_i \geq 0$ is Lagrange genes.

Differentiating L with respect to w , b and ξ_i , and setting the result to zero. The optimization problem (2) can translate into the following simple dual problem.

Maximize:

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

Subject to

$$\sum_{i=1}^l \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C, \quad i = 1, 2, \dots, l$$

Where (x_i, x_j) ($\phi(x_i), \phi(x_j)$) is a kernel function and satisfies the Mercer theorem.

Let α^* is the optimal solutions of (4) and corresponding weight and the bias w^* , b^* , respectively.

According to the Karush-Kuhn-Tucker(KKT) conditions, the solution of the optimal problem (4) must satisfy

$$\alpha_i^* [(x_i \cdot w) + b^* - 1 + \xi_i^*] = 0 \quad i = 1, 2, \dots, l$$

Where α_i^* are non zero only for a subset of vector x_i called support vectors.

Finally, the optimal decision function is

$$f(x) = \text{Sgn} \left(\sum_{i=1}^l y_i \alpha_i^* K(x_i \cdot x) + b^* \right)$$

Where $w^* = \sum_{i=1}^l \alpha_i^* y_i x_i$

1. Fuzzy Support Vector Machines ([12]-[13]).

Consider the aforementioned binary training set S . We choose a proper membership function and receive s_i which is the fuzzy memberships value of the training point x_i . Then, the training set S become fuzzy training set S'

$$S' = \{(x_1, y_1, s_1), (x_2, y_2, s_2), \dots, (x_l, y_l, s_l)\}$$

where $x_i \in \mathbb{R}^n$ and corresponding binary class labels $y_i \in \{1, -1\}$, $0 \leq s_i \leq 1$.

Then, the quadratic programming problem for classification can be described as following:

Minimize:

$$\phi(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l s_i \xi_i$$

Subject to:

$$y_i [(w \cdot x_i) + b] \geq 1 - \xi_i$$

$$\xi_i \geq 0, \quad i = 1, 2, \dots, l$$

where $C > 0$ is punishment gene and ξ_i is a slack variable. The fuzzy membership s_i is attitude of the corresponding point x_i toward one class. It shows that a smaller s_i can reduce the effect of the parameter ξ_i in problem (18), so the corresponding point x_i can be treated as less important

By using the KKT condition and Lagrange Multipliers. We are able to form the following equivalent dual problem

Maximize:

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

Subject to:

$$\sum_{i=1}^l \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq s_i C, \quad i = 1, 2, \dots, l$$

If $\alpha_i > 0$, then the corresponding point x_i is support vectors. More, if $0 < \alpha_i < s_i C$, then support vectors x_i lies round of separating surface; if $\alpha_i = s_i C$, then support vectors x_i belongs to error sample. Then, the decision function of the corresponding optimal separating surface becomes

$$f(x) = \text{Sgn} \left(\sum_{x_i \in S_0} \alpha_i y_i K(x_i \cdot x) + b \right)$$

Where $K(x_i, x)$ is kernel function.

VII. EXPERIMENT

We will investigate the performance of these two techniques, (1) Classifying FSVM on original matrix where Vector Space Model is used, (2) and FSVM on a matrix where Random Indexing is used to reduce the dimensionality of the matrix before singular value decomposition. Performance will be measured as calculation time as well as precision and recall. We have used a subset of the Reuters-21578 text corpus. The subset comprises 3299 that include 5 most frequent categories : earn, acquisition, money, grain, crude.

No	Classifier	F-score	
		FSVM	FSVM+LSI+RI
1	Earn	0.97	0.993

2	Acquisition	0.93	0.965
3	Money	0.95	0.972
4	Grain	0.79	0.933
5	Crude	0.92	0.961
Average		0.912	0.9648

Table 1: The experiment results of FSVM and FSVM+LSI+RI classifiers.

VIII. CONCLUSION

This article introduces Fuzzy Support Vector Machines for Text Categorization based on reduced matrices use Latin Semantic Indexing combined with Random Indexing. Our results show that FSVM_LSI_RI provide better results on Precision and Recall than FSVM. Due to time limit, only experiments on the 5 categories. Future direction include how to use this scheme to future direction include how to use this scheme to classify student's idea at University of Information Technology HoChiMinh City.

REFERENCES

- [1]. April Kontostathis (2007), "Essential Dimensions of latent semantic indexing", Department of Mathematics and Computer Science Ursinus College, Proceedings of the 40th Hawaii International Conference on System Sciences, 2007.
- [2]. Cherukuri Aswani Kumar, Suripeddi Srinivas (2006) , "Latent Semantic Indexing Using Eigenvalue Analysis for Efficient Information Retrieval", Int. J. Appl. Math. Comp. Sci., 2006, Vol. 16, No. 4, pp. 551-558.
- [3]. David A.Hull (1994), *Information retrieval Using Statistical Classification*, Doctor of Philosophy Degree, The University of Stanford.
- [4]. Gabriel Oksa, Martin Becka and Marian Vajtersic (2002), "Parallel SVD Computation in Updating Problems of Latent Semantic Indexing", *Proceeding ALGORITHM 2002 Conference on Scientific Computing*, pp. 113 - 120.
- [5]. Katarina Blom, (1999), *Information Retrieval Using the Singular Value Decomposition and Krylov Subspace*, Department of Mathematics Chalmers University of Technology S-412 Goteborg, Sweden
- [6]. Kevin Erich Heinrich (2007), *Automated Gene Classification using Nonnegative Matrix Factorization on Biomedical Literature*, Doctor of Philosophy Degree, The University of Tennessee, Knoxville.
- [7]. Miles Efron (2003). Eigenvalue - Based Estimators for Optimal Dimensionality Reduction in Information Retrieval. *ProQuest Information and Learning Company*.
- [8]. Michael W. Berry, Zlatko Drmac, Elizabeth R. Jessup (1999), "Matrix, Vector Space, and Information Retrieval", *SIAM REVIEW* Vol 41, No. 2, pp. 335 - 352.
- [9]. Nordinah Ab Samat, Masrah Azrifah Azmi Murad, Muhamad Taufik Abdullah, Rodziah Atan (2008), "Term Weighting Schemes Experiment Based on SVD for Malay Text Retrieval", Faculty of Computer Science and Information Technology University Putra Malaysia, IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.10, October 2008.
- [10]. Jussi Karlgren and Magnus Sahlgren. 2001. From words to understanding. In Y. Uesaka, P.Kanerva, and H. Asoh, editors, *Foundations of Real-World Intelligence*, chapter 26, pages 294-308. Stanford: CSLI Publications.
- [11]. Magnus Rosell, Martin Hassel, Viggo Kann: "Global Evaluation of Random Indexing through Swedish Word Clustering Compared to the People's Dictionary of Synonyms", (Rosell et al., 2009).

- [12]. Shigeo Abe and Takuya Inoue (2002), "Fuzzy Support Vector Machines for Multiclass Problems", *ESANN'2002 proceedings*, pp. 113-118.
- [13]. Shigeo Abe and Takuya Inoue (2001), "Fuzzy Support Vector Machines for Pattern Classification", *In Proceeding of International Joint Conference on Neural Networks (IJCNN '01)*, volume 2, pp. 1449-1454.
- [14]. T.Joachims (1998), "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," *in Proceedings of ECML-98, 10th European Conference on Machine Learning*, number 1398, pp. 137–142.

AUTHORS PROFILE

The author born in 1969 in Da Nang, VietNam. He graduated University of Odessa (USSR), in 1992, specialized in Information Technology. He postgraduated on doctoral thesis in 1996 at the Academy of Science of Russia, specialized in IT. Now he is the Dean of Software Engineering of University of Information Technology, VietNam National University HoChiMinh City.

Research: Knowledge Engineering, Information Systems and software Engineering.

CATEGORIES OF UNSTRUCTURED DATA PROCESSING AND THEIR ENHANCEMENT

Prof.(Dr). Vinodani Katiyar

Sagar Institute of Technology and Management Barabanki U.P. (INDIA)

Hemant Kumar Singh

Azad Institute of Engineering & Technology Lucknow, U.P. INDIA.

ABSTRACT

Web Mining is an area of Data Mining which deals with the extraction of interesting knowledge from the World Wide Web. The central goal of the paper is to provide past, current evaluation and update in each of the three different types of web mining i.e. web content mining, web structure mining and web usages mining and also outlines key future research directions.

Keywords: Web mining; web content mining; web usage mining; web structure mining;

1. INTRODUCTION

The amount of data kept in computer files and data bases is growing at a phenomenal rate. At the same time users of these data are expecting more sophisticated information from them. A marketing manager is no longer satisfied with the simple listing of marketing contacts but wants detailed information about customers' past purchases as well as prediction of future purchases. Simple structured / query language queries are not adequate to support increased demands for information. Data mining steps is to solve these needs. Data mining is defined as finding hidden information in a database alternatively it has been called exploratory data analysis, data driven discovery, and deductive learning [7]. In the data mining communities, there are three types of mining: data mining, web mining, and text mining. There are many challenging problems [1] in data/web/text mining research. Data mining mainly deals with structured data organized in a database while text mining mainly handles unstructured data/text. Web mining lies in between and copes with semi-structured data and/or unstructured data. Web mining calls for creative use of data mining and/or text mining techniques and its distinctive approaches. Mining the web data is one of the most challenging tasks for the data mining and data management

scholars because there are huge heterogeneous, less structured data available on the web and we can easily get overwhelmed with data [2].

According to Oren Etzioni[6] Web mining is the use of data mining techniques to automatically discover and extract information from World Wide Web documents and service. Web mining research can be classified in to three categories: Web content mining (WCM), Web structure mining (WSM), and Web usage mining (WUM) [3]. Web content mining refers to the discovery of useful information from web contents, including text, image, audio, video, etc. Web structure mining tries to discover the model underlying the link structures of the web. Model is based on the topology of hyperlinks with or without description of links. This model can be used to categorize web pages and is useful to generate information such as similarity and relationship between different websites. Web usage mining refers discovery of user access patterns from Web servers. Web usages data include data from web server access logs, proxy server logs, browser logs, user profiles, registration data, user session or transactions, cookies, user queries, bookmark data, mouse clicks and scrolls or any other data as result of interaction.

Minos N. Garofalakis, Rajeev Rastogi, et al[4] presents a survey of web mining research [1999] and analyses Today's search tools are plagued by the following four problems:

(1) The abundance problem, that is, the phenomenon of hundreds of irrelevant documents being returned in response to a search query, (2) limited coverage of the Web (3) a limited query interface that is based on syntactic keyword-oriented search (4) limited customization to individual users

and listed research issues that still remain to be addressed in the area of Web Mining .

Bin Wang, Zhijing Liu[5] presents a survey [2003] of web mining research. With the explosive growth of information sources available on the World Wide Web, it has become more and more necessary for users to utilize automated tools in order to find, extract, filter, and evaluate the desired information and resources. In addition, with the transformation of the web into the primary tool for electronic commerce, it is essential for organizations and companies, who have invested millions in Internet and Intranet technologies, to track and analyze user access patterns. These factors give rise to the necessity of creating server-side and client-side intelligent systems that can

effectively mine for knowledge both across the Internet and in particular web localities. The purpose of the paper is to provide past, current evaluation and update in each of the three different types of web mining i.e. web content mining, web structure mining and web usages mining and also outlines key future research directions.

2. LITERATURE REVIEW

Both Etzioni[6] and Kosala and Blockeel[3] decompose web mining into four subtasks that respectively, are (a) resource finding; (b) information selection and preprocessing; (c) generalization; and (d) analysis. Qingyu Zhang and Richard S. Segall[2] divided the web mining process into the following five subtasks:

- (1) Resource finding and retrieving;
- (2) Information selection and preprocessing;
- (3) Patterns analysis and recognition;
- (4) Validation and interpretation;
- (5) Visualization

The literature in this paper is classified into the three types of web mining: web content mining, web usage mining, and web structure mining. We put the literature into five sections: (2.1) Literature review for web content mining; (2.2) Literature review for web usage mining; (2.3) Literature review for web structure mining; (2.4) Literature review for web mining survey; and (2.5) Literature review for semantic web.

2.1 Web Content Mining- Margaret H. Dunham[7] stated Web Content Mining can be thought of as extending the work performed by basic search engines. Web content mining analyzes the content of Web resources. Recent advances in multimedia data mining promise to widen access also to image, sound, video, etc. content of Web resources. The primary Web resources that are mined in Web content mining are individual pages. Information Retrieval is one of the research areas that provides a range of popular and effective, mostly statistical methods for Web content mining. They can be used to group, categorize, analyze, and retrieve documents. content mining methods which will be used for Ontology learning, mapping and merging ontologies, and instance learning [8].

To reduce the gap between low-level image features used to index images and high-level semantic contents of images in content-based image retrieval (CBIR) systems or search engines, Zhang et al.[9] suggest applying relevance feedback to refine the query or similarity measures in image search process. They present a framework of relevance feedback and semantic learning where low-level features and keyword explanation are integrated in image retrieval and in feedback processes to improve the retrieval performance. They developed a prototype system performing better than traditional approaches.

The dynamic nature and size of the Internet can result in difficulty finding relevant information. Most users typically express their information need via short queries to search engines and they often have to physically sift through the search results based on relevance ranking set by the search engines, making the process of relevance judgement time-consuming. Chen et al.[10] describe a novel representation technique which makes use of the Web structure together with summarization techniques to better represent knowledge in actual Web Documents. They named the proposed technique as Semantic Virtual Document (SVD). The proposed SVD can be used together with a suitable clustering algorithm to achieve an automatic content-based categorization of similar Web Documents. This technique allows an automatic content-based categorization of web documents as well as a tree-like

graphical user interface for browsing post retrieval document browsing enhances the relevance judgment process for Internet users. They also introduce cluster-biased automatic query expansion technique to interpret short queries accurately. They present a prototype of Intelligent Search and Review of Cluster Hierarchy (iSEARCH) for web content mining.

Typically, search engines are low precision in response to a query, retrieving lots of useless web pages, and missing some other important ones. Ricardo Campos et al[11] study the problem of the hierarchical clustering of web and proposed an architecture of a meta-search engine called WISE that automatically builds clusters of related web pages embodying one meaning of the query. These clusters are then hierarchically organized and labeled with a phrase representing the key concept of the cluster and the corresponding web documents.

Mining search engine query log is a new method for evaluating web site link structure and information architecture. Mehdi Hosseini, Hassan Abol hassani [12] propose a new query-URL co-clustering for a web site useful to evaluate information architecture and link structure. Firstly, all queries and clicked URLs corresponding to particular web site are collected from a query log as bipartite graph, one side for queries and the other side for URLs. Then a new content free clustering is applied to cluster queries and URLs concurrently. Afterwards, based on information entropy, clusters of URLs and queries will be used for evaluating link structure and information architecture respectively.

Data available on web is classified as structured data, semi structured data and Unstructured data. Kshitija Pol, Nita Patil et al[13] presented a survey on web content mining described various problems of web content mining and techniques to mine the Web pages including structured and semi structured data.

2.2 Web Structure Mining-Web information retrieval tools make use of only the text on pages, ignoring valuable information contained in links. Web structure mining aims to generate structural summary about web sites and web pages. The focus of structure mining is on link information

[14].Through an original algorithm for hyperlink analysis called HITS (Hypertext Induced Topic Search), Kleinberg[15] introduced the concepts of hubs (pages that refer to many pages) and authorities (pages that are referred by many pages)[16]. Apart from search ranking, hyperlinks are also useful for finding Web communities. A web community is a collection of web pages that are focused on a particular topic or theme. Most community mining approaches are based on the assumption that each member of a community has more hyperlinks within than outside its community. In this context, many graph clustering algorithms may be used for mining the community structure of a graph as they adopt the same assumption, i.e. they assume that a cluster is a vertex subset such that for all of its vertices, the number of links connecting a vertex to its cluster is higher than the number of links connecting the vertex outside its cluster[17].

Furnkranz[18] described the Web may be viewed as a (directed) graph whose nodes are the documents and the edges are the hyperlinks between them and exploited the graph structure of the World Wide Web for improved retrieval performance and classification accuracy. Many search engines use graph properties in ranking their query results.

The continuous growth in the size and use of the Internet is creating difficulties in the search for information. To help users search for information and organize information layout, Smith and Ng[19] suggest using a SOM to mine web data and provide a visual tool to assist user navigation. Based on the users' navigation behavior, they develop LOGSOM, a system that utilizes SOM to organize web pages into a two-dimensional map. The map provides a meaningful navigation tool and serves as a visual tool to better understand the structure of the web site and navigation behaviors of web users.

As the size and complexity of websites expands dramatically, it has become increasingly challenging to design websites on which web surfers can easily find the information they seek.

Fang and Sheng[20] address the design of the portal page of a web site. They try to maximize the efficiency, effectiveness, and usage of a web site's portal page by selecting a limited number of hyperlinks from a large set for the inclusion in a

portal page. Based on relationships among hyperlinks (i.e. structural relationships that can be extracted from a web site and access relationship that can be discovered from a web log), they propose a heuristic approach to hyperlink selection called Link Selector.

Instead of clustering user navigation patterns by means of a Euclidean distance measure, Hay et al.[21] use the Sequence Alignment Method (SAM) to partition users into clusters, according to the order in which web pages are requested and the different lengths of clustering sequences. They validate SAM by means of user traffic data of two different web sites and results show that SAM identifies sequences with similar behavioral patterns.

To meet the need for an evolving and organized method to store references to web objects, Guan and McMullen³⁷ design a new bookmark structure that allows individuals or groups to access the bookmark from anywhere on the Internet using a Java-enabled web browser. They propose a prototype to include more features such as URL, the document type, the document title, keywords, date added, date last visited, and date last modified as they share bookmarks among groups of users.

Song and Shepperd[22] view the topology of a web site as a directed graph and mine web browsing patterns for e-commerce. They use vector analysis and fuzzy set theory to cluster users and URLs. Their frequent access path identification algorithm is not based on sequence mining.

2.3 Web Usages Mining- Several surveys on Web usage mining exist in [3, 23, 24] Web usage mining model is a kind of mining to server logs. And its aim is getting useful users' access information in logs to make sites can perfect themselves with pertinence, serve users better and get more economy benefit The main areas of research in this domain are Web log data preprocessing and identification of useful patterns from this preprocessed data using mining techniques. Most data used for mining [23] is collected from Web servers, clients, proxy servers, or server databases, all of which generate noisy data. Because Web mining is sensitive to noise, data cleaning methods are necessary. Jaideep Srivastava and R. Cooley [23] categorize data preprocessing into subtasks

and note that the final outcome of preprocessing should be data that allows identification of a particular user's browsing pattern in the form of page views, sessions, and click streams. Click streams are of particular interest because they allow reconstruction of user navigational patterns In the previous six years collection of user navigation session were presented in form of many models such as Hyper Text Probabilistic Grammar (HPG), N-Gram Model, Dynamic clustering based morkov model etc[25].Using a footstep graph, The user's click stream data can be visualized and any interesting pattern can be discovered more easily and quickly than with other visualization tools. Recent work by Yannis Manolopoulos, A Nanopoulos et al[26] provides a comprehensive discussion of Web logs for usage mining and suggests novel ideas for Web log indexing. Such preprocessed data enables various mining techniques.

Recently, several Web Usage Mining algorithms [27, 28, 29] have been proposed to mining user navigation behavior. Partitioning method was one of the earliest clustering methods to be used in Web usage mining [28].Web based recommender systems are very helpful in directing the users to the target pages in particular web sites. Web usage mining recommender systems have been proposed to predict user's intention and their navigation behaviors. We can take into account the semantic knowledge [explained in later section] about underlying domain to improve the quality of the recommendation. Integrating semantic web and web usage mining can achieve best recommendations in the dynamic huge web sites [30].

As new data is published every day, the Web's utility as an information source will continue to grow. The only question is: Can Web mining catch up to the WWW's growth? There are existing Web Usages mining models for modeling the user navigation patterns. My work will be an effort to advance the existing web usages mining system and to present the work principle of the system. The key technologies in system design are session identification, data cleaning and web Personalization.

2.4 Web Mining- In 1996 it's Etzioni [6] who first coined the term web mining.Etzioni starts by making a hypothesis that

information on web is sufficiently structured and outlines the subtasks of web mining and describes the web mining process. Web mining may be decomposed into the following sub tasks:

1. **Resource Discovery:** locating unfamiliar documents and services on the Web.
2. **Information Extraction:** automatically extracting specific information from newly discovered Web resources.
3. **Generalization:** uncovering general patterns at individual Web sites and across multiple Sites.

Kosala and Blockeel[3] who perform research in the area of web mining and suggest the three web mining categories of web content, web structure, and web usage mining.

Han and Chang[32] author a paper on data mining for web intelligence that claims that “incorporating data semantics could substantially enhance the quality of keyword-based searches,” and indicate research problems that must be solved to use data mining effectively in developing web intelligence. The latter includes mining web search-engine data and analyzing web’s link structure, classifying web documents automatically, mining web page semantic structures and page contents, and mining web dynamics. Web dynamics is the study of how the web changes in the context of its contents, structure, and access patterns.

Barsagade[33] provides a survey paper on web mining usage and pattern discovery. Chau et al.[34] discuss personalized multilingual web content mining. Kolari and Joshi [35] provide an overview of past and current work in the three main areas of web mining research-content, structure, and usage as well as emerging work in semantic web mining.

Scime⁴⁵ edit a “Special Issue on Web Content Mining” of the Journal of Intelligent Information Systems (JIIS).

2.5 Semantic Web Mining- The Semantic Web[37] is a web that is able to describe things in a way that computers can understand. Statements are built with syntax rules. The syntax of a language defines the rules for building the language statements. But how can syntax become semantic? This is what the Semantic Web is all about. Describing things in a way that computers applications can understand it. The Semantic Web is not about links between web pages. The Semantic Web describes the relationships between things (like

A is a part of B and Y is a member of Z) and the properties of things (like size, weight, age, and price). Semantic Web Mining aims at combining the two fast-developing research areas Semantic Web and Web Mining. More and more researchers are working on improving the results of Web Mining by exploiting semantic structures in the Web, and they make use of Web Mining techniques for building the Semantic Web. Last but not least, these techniques can be used for mining the Semantic Web itself [38]. The Semantic Web is a recent initiative, inspired by Tim Berners-Lee[39], to take the World-Wide Web much further and develop in into a distributed system for knowledge representation and computing. The aim of the Semantic Web is to not only support access to information “on the Web” by direct links or by search engines but also to support its use. Instead of searching for a document that matches keywords, it should be possible to combine information to answer questions. Instead of retrieving a plan for a trip to Hawaii, it should be possible to automatically construct a travel plan that satisfies certain goals and uses opportunities that arise dynamically. This gives rise to a wide range of challenges. Some of them concern the infrastructure, including the interoperability of systems and the languages for the exchange of information rather than data. Many challenges are in the area of knowledge representation, discovery and engineering. They include the extraction of knowledge from data and its representation in a form understandable by arbitrary parties, the intelligent questioning and the delivery of answers to problems as opposed to conventional queries and the exploitation of formerly extracted knowledge in this process .

3.0 CONCLUSION-

This paper has provided a more current evaluation and update of web mining research available. Extensive literature has been reviewed based on three types of web mining, namely web content mining, web usage mining, and web structure mining. This paper helps researchers and practitioners effectively accumulate the knowledge in the field of web mining, and speed its further development.

4.0 FUTURE RESEARCH DIRECTIONS-

1. Investigation into Semantic Web applications such as that for bioinformatics in which biological data and knowledge bases are interconnected.
2. Applications of intelligent personal assistant or intelligent software agent that automatically accumulates and classifies suitable information based on user preference
3. Although we have focused on representing knowledge in HTML Web Documents, there are numerous other file formats that are publicly accessible on the Internet. Also, if both the actual Web Documents and corresponding Back Link Documents were mainly composed of multimedia information (e.g. graphics, audio, etc.), SVD will not be particularly effective in revealing more textual information. It would be worthwhile to research new techniques to include these file formats and multimedia information for knowledge representation.

REFERENCE

- [1] Q. Yang and X. Wu, 10 challenging problems in data mining research, *Int. J Inform. Technol. Decision Making* 5(4) (2006) 597–604.
- [2] Qingyu Zhang and Richard s. Segall,” Web mining: a survey of current research, Techniques, and software”, in the *International Journal of Information Technology & Decision Making* Vol. 7, No. 4 (2008) 683–720
- [3] Kosala and Blockeel, “Web mining research: A survey,” *SIGKDD:SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery and Data Mining*, ACM, Vol. 2, 2000
- [4] Minos N. Garofalakis, Rajeev Rastogi, et al “Data Mining and the Web: Past, Present and Future” *Proceedings of the 2nd international workshop on Web information and data management* Kansas City, Missouri, United States pp: 43 - 47 (1999)
- [5] Bin Wang, Zhijing Liu, "Web Mining Research," *iccima*, pp.84, Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'03), 2003
- [6] O.etzioni. The world wide web: Quagmire or Gold Mining. *Communicate of the ACM*, (39)11:65-68, 1996;
- [7] Margaret H. Dunham, “Data Mining Introductory & Advanced Topics”, Pearson Education
- [8] Semantic Web Mining: State of the art and future directions” *Web Semantics: Science, Services and Agents on the World Wide Web*, Volume 4, Issue 2, June 2006, Pages 124-143
- [9] H. Zhang, Z. Chen, M. Li and Z. Su, Relevance feedback and learning in content-based image search, *World Wide Web* 6(2) (2003) 131–155.
- [10] L. Chen, W. Lian and W. Chue, Using web structure and summarization techniques for web content mining, *Inform. Process. Management: Int. J.* 41(5) (2005) 1225–1242
- [11] Ricardo Campos, Gael Dias, Celia Nunes, "WISE: Hierarchical Soft Clustering of Web Page Search Results Based on Web Content Mining Techniques," *wi*, pp.301-304, 2006 *IEEE/WIC/ACM International Conference on Web Intelligence (WI'06)*, 2006.
- [12] Mehdi Hosseini, Hassan Abolhassani,” Mining Search Engine Query Log for Evaluating Content and Structure of a Web Site”, *International Conference on Web Intelligence* 2007
- [13] Kshitija Pol, Nita Patil et al,” A Survey on Web Content Mining and extraction of Structured and Semistructured data” in *Proceedings of the 2008 First International Conference on Emerging Trends in Engineering and Technology*.
- [14] Sanjay Kumar Madria , Sourav S. Bhowmick , Wee Keong Ng , Ee-Peng Lim, *Research Issues in Web Data Mining*, *Proceedings of the First International Conference on Data Warehousing and Knowledge Discovery*, p.303-312, September 01, 1999
- [15] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.
- [16] Nacim Fateh Chikhi, Bernard Rothenburger, Nathalie Aussenac-Gilles “A Comparison of Dimensionality Reduction Techniques for Web Structure Mining”,

- Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, P.116-119, 2007
- [17] Lefteris Moussiades, Athena Vakali, "Mining the Community Structure of a Web Site," bci, pp.239-244, 2009 Fourth Balkan Conference in Informatics, 2009
- [18] J. Furnkranz, Web structure mining — Exploiting the graph structure of the worldwide web, "OGAI-J. 21(2) (2002) 17–26
- [19] K. A. Smith and A. Ng, Web page clustering using a self-organizing map of user navigation patterns, Decision Support Syst. 35(2) (2003) 245–256
- [20] X. Fang and O. Sheng, LinkSelector: A web mining approach to hyperlink selection for web portals, ACM Trans. Internet Tech. 4(2) (2004) 209–237
- [21] B. Hay, G. Wets and K. Vanhoof, Mining navigation patterns using a sequence alignment method, Knowledge Inform. Syst. 6(2) (2004) 150–163
- [22] Q. Song and M. Shepperd, Mining web browsing patterns for e-commerce, Comput. Indus. 57(7) (2006) 622–630
- [23] Jaideep Srivastava, R. Cooley, "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data", ACM SIGKDD, VOL.7 No. 2 Jan 2000
- [24] Subhash K.Shinde, Dr.U.V.Kulkarni, "A New Approach For On Line Recommender System in Web Usage Mining", Proceedings of the 2008 International Conference on Advanced Computer Theory and Engineering Pages: 973-977
- [25] Borges and M. Levene, "A dynamic clustering-based markov model for web usage Mining", cs.IR/0406032, 2004
- [26] Yannis Manolopoulos et al, "Indexing web access-logs for pattern queries", Workshop On Web Information And Data Management archive Proceedings of the 4th international workshop on Web information and data management pp: 63 - 68 2002
- [27] B. Liu and K. Chang, Editorial: Special issue on web content mining, SIGKDD Explorations 6(2) (2004) 1–4
- [28] M. Jalali, N. Mustapha, M. N. Sulaiman, and A. Mamat, "Web User Navigation Pattern mining approach based on Graph Partitioning algorithm " Journal of Theoretical and Applied Information Technology vol. 4, pp. 1125-1130, 2008
- [29] Zhang Huiying; Liang Wei;,"An intelligent algorithm of data pre-processing in Web usage mining," Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on , vol.4, no., pp. 3119- 3123 Vol.4, 15-19 June 2004
- [30] Mehdi Hosseini , Hassan Abol hassani , "Mining Search Engine Query Log for Evaluating Content and Structure of a Web Site" in Proceedings of the 2007 IEEE/WIC/ACM International Conference on Web Intelligence.
- [31] J. Han and C. Chang, Data mining for web intelligence, Computer (November 2002),pp. 54–60, <http://www-faculty.cs.uiuc.edu/~hanj/pdf/computer02.pdf>
- [32] N. Barsagade, Web usage mining and pattern discovery: A survey paper, Computer Science and Engineering Dept., CSE Tech Report 8331 (Southern Methodist University,Dallas, Texas, USA, 2003).
- [33] R. Chau, C. Yeh and K. Smith, Personalized multilingual web content mining, KES (2004), pp. 155–163
- [34] P. Kolari and A. Joshi, Web mining: Research and practice, Comput. Sci. Eng.July/August (2004) 42–53
- [35] A. Scime, Guest Editor's Introduction: Special Issue on Web Content Mining: Special Issue on Web Content Mining, J. Intell. Inform. Syst. 22(3) (2004) 211–213
- [36] W3Schools, Semantic web tutorial (2008) <http://www.w3schools.com/semweb/default.asp>
- [37] Semantic Web Mining:State of the art and future directions" Web Semantics: Science, Services and Agents on the World Wide Web, Volume 4, Issue 2, June 2006, Pages 124-143
- [38] Berners-Lee, T., Fischetti, M.: Weaving the Web. Harper, San Francisco (1999)
- [39] Bettina Berendt, Andreas Hotho, Dunja Mladenic, Maarten van Someren, Myra Spiliopoulou and Gerd Stumme "A Roadmap for Web Mining:From Web to Semantic Web"DOI: 10.1007/978-3-540-30123-3_1

False Positive Reduction using IDS Alert Correlation Method based on the Apriori Algorithm

Homam El-Taj, Omar Abouabdalla, Ahmed Manasrah,
Mohammed Anbar, Ahmed Al-Madi

National Advanced IPv6 Center of Excellence (NAv6)
Universiti Sains Malaysia

Penang, Malaysia

Abstract—Correlating the Intrusion Detection Systems (IDS) is one challenging topic in the field of network security. There are many benefits from correlating the IDS alerts: to reduce the huge amount of alerts that IDS triggers, to reduce the false positive ratio and to figure out the relations between the alerts to get better understanding of the attacks. One of these correlation techniques based on the data mining. In this paper we developed new IDS alerts group correlation method (GCM) based on the aggregated alerts by the Threshold Aggregation Framework (TAF) we create our correlation method by adapting the *Apriori* algorithm for large data. This method used to reduce the amount of aggregated alerts and to reduce the ratio of false positive alerts.

Keywords—Intrusion Detection System; False Positive Alerts; Alert Correlation; Data Mining.

I. INTRODUCTION

Based on the essential and extensive usage of internet and their applications, threats and intrusions become wider and smarter. And because IDS triggers huge amount of alerts the need of study these alerts become essential too. The study of IDS alerts led to bringing to light some of the IDS issues which should be studied, these issues comes in how to group the alerts, define the relation between the alerts and reduce the false alerts.

II. INTRUSION DETECTION SYSTEM (IDS)

IDS monitors the protected network activities and analyze them to trigger alerts if there is any malicious activity accrued. IDS can detect these activities based on anomaly detection methods [1], misuse detection methods [2] or a compensation between both of them. While anomaly methods detect the malicious traffic by determining the abnormality between the suspicious activities flow and the norm flow based on a chosen threshold, misuse methods detect malicious activities based on their signatures. The main differences between these methods based on the detecting novel attacks and the false positive ratio, misuse

methods have minimum amount of false positive, while anomaly methods can detect novel attacks.

III. IDS ALERTS' CORRELATION STUDIES

Correlation is part of intrusion detection studies that smoothes the progress of the analysis of intrusion alerts based on the similarity between alert attributes, this can be represented in mathematical expression as below:

$$Corr_Alert = \{Alert_1, Alert_2, \dots, Alert_n\}$$

Where the group of alerts $\{Alert_1, Alert_2, \dots, Alert_n\}$ with the same features which have relations is represented by $Corr_Alert$. However, most of the correlation methods focus on IDS alerts by examining other intrusion evidence provided by system monitoring tools or scanning tools. The aim of correlation analysis is to detect relationships among alerts so it will be easy to build attack scenarios.

A. Classification of Alert Correlation Technique

IDS alerts correlation studies got many angles to cover this issue using many methods and techniques which can be categorized by: similarity-based, pre-defined attack scenarios, pre-requisites and consequences and statistical causal analysis.

a) Similarity-Based

This technique is based on comparing alert features to see if there is a similarity between the features, mainly the correlation will be based on these features (Source IPs, Destination IPs, Source Ports and Destination Ports). Valdes and Skinner [3] correlated the IDS alerts by three phases starting with the minimum similarity is based on the similarity of source and destination IPs, while the second phase similarity is based on attack class and attack name plus source and destination IPs. This phase ensures that it correlates the same alert from different sensors, and the last phase a threshold value is applied to correlate two alerts based on the similarity of similar attack class with no consideration of other features.

b) Pre-Defined Attack scenarios

The idea of studying the attack scenarios came from the fact that intrusions mainly took several actions to a successful attack.

Debar and Wespi [4] They proposed a system to correlate and aggregate IDS alerts triggers by different sensors, their system got two steps starting by removing the redundant alerts if they are from different sensor, then correlating the alerts is achieved by applying the Consequences rules which specifies that any alert should be followed by another type of alert, depending on these rules the alerts will be correlated so the aggregation phase will start to check if there are any similarity between the source and destination IPs and attack class.

c) Pre-Requisites and Consequences

This technique comes in the middle between features similarity correlations and scenarios based correlations. Pre-requisites can be defined as the essential conditions that must exist for the attack to be succeeded, and consequences for the attacks are defined as conditions that might exist after a specific attack occurred.

Cuppens and Mieke [5] they proposed a cooperation module for IDS alerts with five main functions: alert base management function to normalize the alerts, alert clustering and alert merging functions used to detect the similarity so the alerts will be clustered and merged with each other, alert correlation function will use the explicit correlation rules with pre-defined and consequence statement to do the correlation, intention recognition function which is used to extrapolate intruder actions provides a global diagnosis of the (past, present and future) of the intruders actions, and reaction function used to help the system administrators to choose the best measurement to prevent the intruder's malicious actions.

d) Statistical Causal Analysis

This technique relies on the way of ranking the IDS alerts based on one of the statistical models to correlate them.

Kumar et.al [6] implemented anomaly detection by using Granger Causality Test (time series analysis method) to correlate alerts in attack scenario analysis. This technique aims to reduce the amount of raw alerts by merging alert based on their features, statistical causal analysis uses clustering technique to rank the alerts based on the relations of attacks. This technique is a pure statistical causality analysis with no need for a pre-defined knowledge attack scenarios.

IV. PROPOSED ALERT CORRELATION METHOD USING THE APRIORI ALGORITHM

Our correlation method is based on the IDS aggregated alert using Threshold Aggregation Framework (TAF), TAF output will be accurate aggregated alerts with no redundant alerts and incomplete alerts. In TAF to aggregate two alerts or more a threshold value should be applied to give more accuracy combination results [7].

Figure 4.1 shows the TAF flowchart, the TAF has two types of inputs; the IDS alerts and the user aggregation options. Depending on these two inputs the aggregation will be done. The user will choose which type of aggregation method to aggregate the IDS alerts.

We propose Group Correlation Method (GCM) which will use the output of the TAF to correlate the alerts by using the *Apriori* algorithm.

From the GCM flowchart in Figure 4.2 we can see that there is an alert counter checker to see whether the amount of the alert in the file less than or equal 2 we drop the alerts since there will be no need to correlate them.

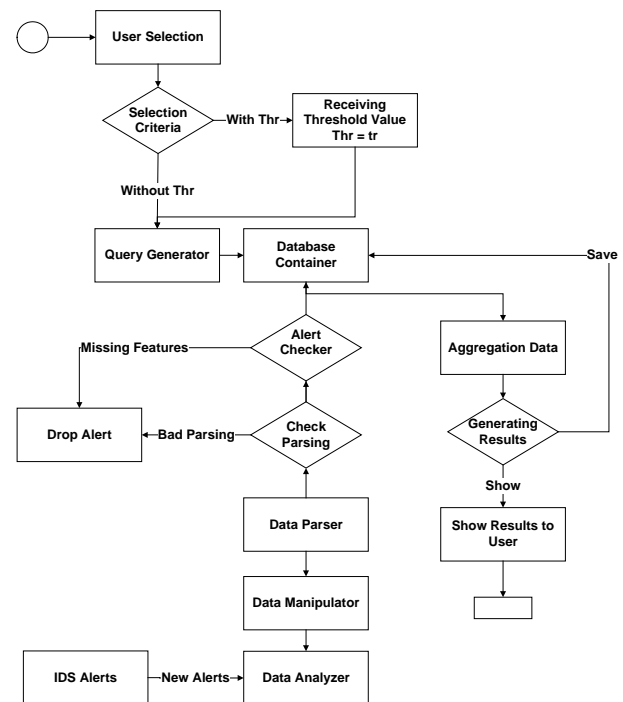


Figure 4.1 TAF flowchart [7]

A. Apriori Algorithm

The reason of choosing the *Apriori* algorithm because it is one of fastest data mining algorithms used to find all frequent itemsets in a large database[8]. *Apriori* algorithm depends on two predefined threshold values (Support and Confidence) to see whether the itemset (group of alerts) are related to each other or not. The Support value equals the frequent of items in the itemset, while the Confidence value can be calculated by the following equation:

$$\text{Confidence} = \frac{\text{LHD} + \text{RHD}}{\text{LHD}} * 100\% \quad (1)$$

Where LHD is the support of left side, RHD is the support of right side.

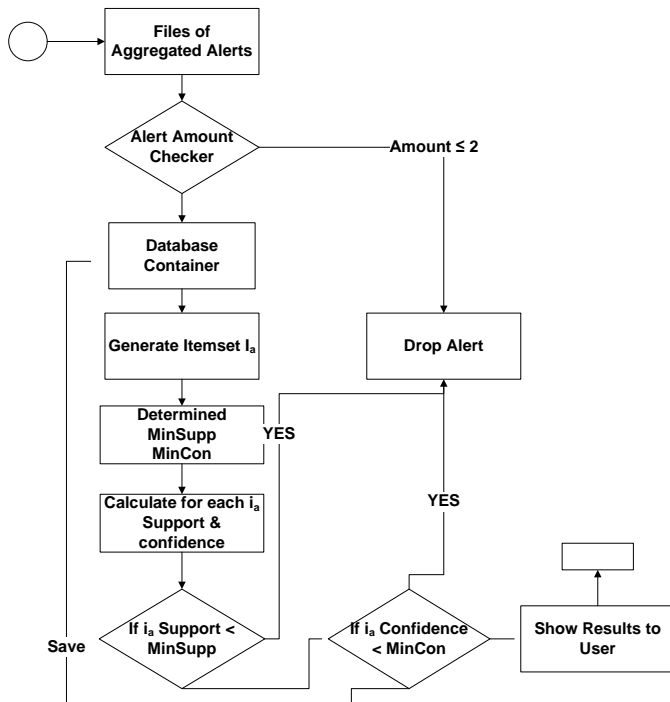


Figure 4.2 GCM flowchart

Support value should be calculated first for each itemset in the current iteration, and only the itemsets that are bigger than the threshold value *minSupp*. The second step is to calculate the confidence by using equation 1. this step will be done for each itemset in the current iteration, this confidences value will be compared with the second threshold value *minCon* to determine whether the current itemset will be used in the second iteration or not. However; the main idea of *Apriori* is to determine if there is a relationship between the alerts which will be distinguished by the confidence value.

Apriori works as illustrated in figure 4.3:

- (1) Read the aggregated alert
- (2) Get two Items as a set of the First Item and the value of the redundant of that

- Item in the second item group as one set of $S\{i_1, i_2, \dots, i_n\}$
- (3) Set *minSupp* & Set *minCon*
- (4) Calculate support value for each i_n in S
- (5) Iteration $I = n-1$
- (6) While $I \geq 1$
- (7) Do $\bigcap_{r=1}^n i_{ar}$
- (8) Calculate Support and Confidence for i_n in $D\{j_1, j_2, \dots, j_m\}$ where $D \in S$
- (9) For each j_m in D if Support < *minSupp* OR Confidence < *minCon* Drop the Itemset.
- (10) $I = I-1$

Figure 4.3 Apriori Algorithm

B. Mathematical representation of Apriori Algorithm

For a better understanding of *Apriori* algorithm we are mathematically representing it as follow:

The Initial Step:-

Let Itemset $S = i_1, i_2, \dots, i_n$, $R = 1, 2, 3, \dots, g$ and $I = \text{Iteration}$.

Iteration $I=0$:-

$S = (i_1, i_2, \dots, i_n)$, $T = (j_1, j_2, \dots, j_k)$ such that $jm \in \{1, 2, 3, \dots, g\}$, $m = (1, 2, \dots, k)$

$\text{Support} = |T| = k$

Iteration $I=1$:-

We make intersection between i_e and i_d where $e \neq d$ such that

$$i_f \cap i_d = (j_1, j_2, \dots, j_k)_e \cap (j_1, j_2, \dots, j_k)_d = (j_1, j_2, \dots, j_b)$$

Where, $j_1, j_2, \dots, j_b \in 1, 2, 3, \dots, g$ and $b \leq e, b \leq d$

Let

$$i_{ed} = i_e \cap i_d$$

$$S = i_{ed}$$

Where, $e = 1, \dots, n$, and $d = 1, \dots, n$

$e \neq d$

$$T = i_e \cap i_d$$

$$\text{Support} = |T| = b$$

If $b < minSupp$ then eliminate i_{ed}

Iteration $I=2$:-

We make intersection between Three $S i_e, i_d$ & i_h

$$i_f \cap i_d \cap i_h = (i_f \cap i_d)_e \cap i_h = (j_1, j_2, \dots, j_b) \cap (j_1, j_2, \dots, j_k) = (j_1, j_2, \dots, j_f)$$

Where, $j_1, j_2, \dots, j_f \in 1, 2, 3, \dots, g$ and $f \leq b, f \leq h$

$$S = i_{edh}$$

Where, $e = 1, \dots, n$ and $d = 1, \dots, n$ and $h = 1, \dots, n$

$$e \neq d \neq h$$

$$T = i_e \cap i_d \cap i_h$$

$$Support = |T| = f$$

If $b < minSupp$ then eliminate i_{edh}

Iteration $I=c$:- (General Form)

We make intersection between each itemset in c $S = i_{a1}, i_{a2}, \dots, i_{ac}$

$$i_{a1} \cap i_{a2} \dots \cap i_{ac} = \bigcap_{r=1}^c i_{ar} = (j_1, j_2, \dots, j_z)$$

Where, $j_1, j_2, \dots, j_z \in 1, 2, 3, \dots, g$ and $z \leq$ from all order in S

$$S = \bigcap_{r=1}^c i_{ar}$$

$$Support = |T| = z \quad \text{If } b < minSupp \quad \text{then} \quad \text{eliminate} \quad \bigcap_{r=1}^c i_{ar}$$

$$Confidence \text{ of } S = \frac{z}{Support \bigcap_{r=1}^c i_{ar}}$$

Remark:

The denominator

$$(Support \bigcap_{r=1}^c i_{ar})$$

representing the Support of all components in

$$\bigcap_{r=1}^c i_{ar}$$

The confidence should be calculated for each itemset, and

then the average of all confidence for that itemset will be the confidence for it.

To understand the mathematical representation, check the following Example:

Let the sample of the first item and the second item took from the table 4.2, $minSupp = 2$, $minCon = 80\%$.

TABLE 4.2 EXAMPLE SET

First Item	Second Item
1	1
2	1
5	2
2	3
3	1
4	2
1	2
2	3
3	2
5	2

So First item $F = \{1, 2, 5, 2, 3, 4, 1, 2, 3, 5\}$, and Second Item $S = \{1, 2, 3\}$

$I=0$

$F_0 = \{1, 2, 3, 4, 5\}$ and $S_0 = \{\{1, 2\}, \{1, 2, 3\}, \{1, 2\}, \{2\}, \{2\}\}$ (No redundancy in second Item)

$Support_{F_0} = \{2, 3, 2, 1, 1\}$ (Items (4, 5) will be eliminated $< minSupp$)

$I=1$

$F_1 = \{(1, 2), (1, 3), (2, 3)\}$ and $S_1 = \{\{1, 2\}, \{1, 2\}, \{2\}\}$

$Support_{F_1} = \{2, 2, 1\}$ (Item (2, 3) will be eliminated $< minSupp$)

$$Confidence_{F(1,2)} = Average \left\{ \frac{Support_{(1,2)}}{Support_1} * 100\% + \frac{Support_{(1,2)}}{Support_2} * 100\% \right\} = \frac{100 + 67}{2} = 83\%$$

$$Confidence_{F(1,3)} = Average \left\{ \frac{2}{2} * 100\% + \frac{2}{2} * 100\% \right\} = 100\%$$

$$Confidence_{F(2,3)} = Average \left\{ \frac{1}{3} * 100\% + \frac{1}{2} * 100\% \right\} = \frac{33+50}{2} = 42\% \text{ (Item will be eliminated)}$$

$I=2$

$F_2 = \{(1, 2, 3)\}$ and $S_1 = \{\{1, 2\}\}$

$Support_{F_0} = \{3\}$

$$Confidence_{F(1,2)} = Average \left\{ \frac{2}{2} * 100\% + \frac{2}{2} * 100\% + 0 * 100\% \right\} = 100\%$$

From the above example it is Obvious that: First; the stopping rule of the iterations when there are no items to compare with. Second; the itemsets (1,2), (1,3), (1, 2, 3)

have relationships by their percentage of {83%, 100%, 100%}. Third; the items (4) and (5) are out of range.

V. IMPLEMENTATION ISSUES

Group correlation Method (GCM) can be used as standalone system to read the aggregated IDS alerts, moreover; GCM can work only with complete alerts with no redundancy to correlate them easing the analyst job. GCM has two main inputs: the user to choose his threshold values (*minSupp* and *minCon*) and the aggregated IDS alerts to be correlated. Figure 4.2 shows the GCM flowchart. GCM will start processing the IDS alerts with no need of filtering the alerts or remove the redundancy. Finally the result will be shown and save in the database based on user request. The process of dropping the insufficient alerts means that these alerts have no relationships with other alerts.

VI. DISCUSSION

This paper presented the GCM method for correlating the aggregated alerts from TAF. The advantages of the proposed method are the improvement of the alert correlation process, especially when it is related to accurate irredundant alerts only, and reducing the time for correlating the alerts. The main objective is to minimize the amount of alerts by investigating the relationships between the alerts and alerts' features which will lead to minimizing the false positive form the IDS alerts.

This method intends to become a general guide that can be implemented and extended to full Forensic investigation system. Other benefits of the proposed methods are: Firstly, to discover the attacks' behaviors. Secondly, finding novel attacks. Thirdly, this method will save the time of analyzing the alerts. Finally, using this method will give us relational accurate alerts with no false alerts. Modifying the value of the two thresholds will control the amount of correlated alerts.

ACKNOWLEDGMENT

This research was supported by the National Advanced IPv6 Center of Excellence (NAv6) in Universiti Sains Malaysia (USM).

REFERENCES

- [1] W. Fan, M. Miller, S. Stolfo, W. Lee, and P. Chan, "Using artificial anomalies to detect unknown and known network intrusions," *Knowledge and Information Systems*, vol. 6, pp. 507-527, 2004.
- [2] M. Sheikhan and Z. Jadidi, "Misuse Detection Using Hybrid of Association Rule Mining and Connectionist Modeling," *World Applied Sciences I*, vol. 7, pp. 31-37, 2009.
- [3] A. Valdes and K. Skinner, "Probabilistic alert correlation," in *the Fourth International Symposium on Recent Advances in Intrusion Detection*, 2001, pp. 54-68.
- [4] H. Debar and A. Wespi, "Aggregation and correlation of intrusion-detection alerts," in *4th International Symposium on Recent Advance in Intrusion Detection(RAID) 2001*, 2001, pp. 85-103.

- [5] F. Cuppens and A. Mieke, "Alert correlation in a cooperative intrusion detection framework," in *IEEE Symposium on Security and Privacy*, Berkeley, California, USA, 2002, pp. 202-215.
- [6] V. Kumar, J. Srivastava, A. Lazarevic, W. Lee, and X. Qin, "Statistical Causality Analysis of Infosec Alert Data," in *Managing Cyber Threats*. vol. 5: Springer US, 2005, pp. 101-127.
- [7] Homam El-Taj, Omar Abouabdalla, Ahmed Manasrah, Ahmed Al-Madi, Muhammad Imran Sarwar, and S. Ramadass, "Forthcoming Aggregating Intrusion Detection System Alerts Framework," in *The Fourth International Conference on Emerging Security Information, Systems and Technologies (SECURWARE 2010)*, Venice/Mestre, Italy 2010.
- [8] W. Kusters and W. Pijls, "Apriori, a depth first implementation," in *Frequent Itemset Mining Implementations Repository (FIMI03)*, 2003.

AUTHORS PROFILE

Homam El-Taj Is a research officer and fellowship holder in National Advanced IPv6 Centre of excellence (NAv6) at Universiti Sains Malaysia (USM). He hold his Bachelor in Computer Science From Philadelphia University Amman Jordan 2003, and a Master degree in computer science from (USM) in the area of Distributed Computing 2006, His master research was on Message Digest Based on Elliptic Curve Concept (MDEC), Currently he is a PhD Candidate in NAv6 at USM, His PhD research area in the field of Network Security, He has published several research articles in Journals and Proceedings.

Dr. Omar Amer Abouabdalla obtained his PhD degree in Computer Sciences from University Science Malaysia (USM) in the year 2004. Presently he is working as a senior lecturer and domain head in the National Advanced IPv6 Centre - USM. He has published more than 50 research articles in Journals and Proceedings (International and National). His current areas of research interest include Multimedia Network, Internet Protocol version 6 (IPv6), and Network Security.

Dr. Ahmed M. Manasrah is a senior lecturer and the Head of iNetmon project as well as the research and innovation of the National Advanced IPv6 Centre of Excellence (NAv6) in Universiti Sains Malaysia. He is also the IMPACT Research Domain Head for Botnet and threat assessment Research. Dr. Ahmed obtained his Bachelor of Computer Science from Mutah University, al Karak, Jordan in 2002. He obtained his Master of Computer Science and doctorate from Universiti Sains Malaysia in 2005 and 2009 respectively. Dr. Ahmed is heavily involved in researches carried by NAv6 centre, such as Network monitoring and Network Security monitoring with filing 3 Patents in Malaysia.

Mohammed Anbar is a research officer in National Advanced IPv6 Centre of Excellence (NAv6) at Universiti Sains Malaysia. His main research area is Network Security and Malware Protection. Anbar has achieved his Masters in information technology from university Utara Malaysia (UUM) in 2009. Currently, he is a PhD candidate in NAv6.

Ahmed Azmi Almadi is a research officer in National Advanced IPv6 Centre of Excellence (NAv6) at Universiti Sains Malaysia. His main research area is Network Security and Malware Protection. Almadi has obtained his Masters in Computer Science from USM in 2007. Currently, he is a PhD candidate and fellowship holder in NAv6. His PhD research focuses on Botnet Detection.

Sector Mean with Individual Cal and Sal Components in Walsh Transform Sectors as Feature Vectors for CBIR

H. B. Kekre

Senior Professor, Computer Engineering
MPSTME,SVKM'S NMIMS University,
Mumbai, INDIA

Dhirendra Mishra

Associate Professor, Computer Engineering
MPSTME, SVKM'S NMIMS University,
Mumbai, INDIA

Abstract- We have introduced a novel idea of conceiving complex Walsh transform for sectorization of transformed components. In this paper we have proposed two different approaches for feature vector generation with consideration of all sal components and all cal components separately. Both these approaches are experimented with the extra components of zero-cal and highest-sal. Two similarity measures such as sum of absolute difference and Euclidean distance are used and results are compared. The cross over point performance of overall average of precision and recall for both approaches on different sector sizes are compared. The individual sector mean of Walsh sectors in all three color planes are considered to design the feature vector. The algorithm proposed here is worked over database of 1055 images spread over 12 different classes. Overall Average precision and recall is calculated for the performance evaluation and comparison of 4, 8, 12 & 16 Walsh sectors. The use of Absolute difference as similarity measure always gives lesser computational complexity and consideration of only all cal components with augmentation of zero-cal approach with sum of absolute difference as similarity measure of feature vector has the best retrieval performance.

Index Terms- CBIR, Walsh Transform, Euclidian Distance, Absolute Difference, Precision, Recall

I. INTRODUCTION

With the huge growth of digital information the need of its management requires need of storage and utilization in efficient manner. This has lead to approach like content based image search and retrieval to be used. Content-based image retrieval into automatic retrieval of images from a database by color, texture and shape features. The term has been widely used to describe the process of retrieving desired image on the basis of features (such as colors, texture and shape) that can be automatically extracted from the images themselves. The typical CBIR system [1-6] performs two major tasks. The first one is feature extraction (FE), where a set of features, called image signature or feature vector, is generated to accurately represent the content of each image in the database. A feature vector is

much smaller in size than the original image, typically of the order of hundreds of elements (rather than millions). The second task is similarity measurement (SM), where a distance between the query image and each image in the database using their signatures is computed so that the top closest images can be retrieved.[7-9]. There are various approaches which have been experimented to generate the efficient algorithm for CBIR like FFT sectors [4-6], Transforms [15][17], Vector quantization[12], bit truncation coding [13][14]. In this paper we have introduced a novel concept of complex Walsh transform and its sectorization for feature extraction (FE).Two different similarity measures namely sum of absolute difference and Euclidean distance are considered. The performances of these approaches are compared.

II. WALSH TRANSFORM

Walsh transform [17] matrix is defined as a set of N rows, denoted W_j , for $j = 0, 1, \dots, N - 1$, which have the following properties:

- W_j takes on the values $+1$ and -1 .
- $W_j[0] = 1$ for all j .
- $W_j \times W_k^T = 0$, for $j \neq k$ and $W_j \times W_k^T = N$, for $j=k$.
- W_j has exactly j zero crossings, for $j = 0, 1, \dots, N-1$.
- Each row W_j is either even or odd with respect to its midpoint.

Walsh transform matrix is generated using a Hadamard matrix of order N . The Walsh transform matrix row is the row of the Hadamard matrix specified by the Walsh code index, which must be an integer in the range $[0, \dots, N - 1]$. For the Walsh code index equal to an integer j , the respective Hadamard output code has exactly j zero crossings, for $j = 0, 1, \dots, N - 1$.

Kekre's Algorithm[10] to generate Walsh Transform from Hadamard matrix [17] is illustrated for $N=16$. However the algorithm is general and can be used for any $N = 2^k$ where k is an integer.

Step 1:

Arrange the 'N' numbers in a row and then split the row at 'N/2', the other part is written below the upper row but in reverse order as follows:

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
15 14 13 12 11 10 9 8

Step 2:

We get two rows, each of this row is again split in 'N/4' and other part is written in reverse order below the upper rows as shown below.

0 1 2 3
15 14 13 12
7 6 5 4
8 9 10 11

This step is repeated until we get a single column which gives the ordering of the Hadamard rows according to sequency as given below:

0, 15, 7, 8, 3, 12, 4, 11, 1, 14, 6, 9, 2, 13, 5, 10

Step 3:

According to this sequence the Hadamard rows are arranged to get Walsh transform matrix. Now a product of Walsh matrix and the image matrix is calculated. This matrix contains Walsh transform of all the columns of the given image.

Since Walsh matrix has the entries either +1 or -1 there is no multiplication involved in computing this matrix. Since only additions are involved computational complexity is very low.

III. FEATURE VECTOR GENERATION

The proposed algorithm makes novel use of Walsh transform to design the sectors to generate the feature vectors for the purpose of search and retrieval of database images. The complex Walsh transform is conceived by multiplying all sal functions by $j = \sqrt{-1}$ and combining them with real cal functions of the same sequency. Thus it is possible to calculate the angle by taking \tan^{-1} of sal/cal. However the values of tan are periodic with the period of π radians hence it can resolve these values in only two sectors. To get the angle in the range of 0-360 degrees we divide these points in four sectors as explained below. These four

sectors are further divided into 8, 12 and 16 sectors. We have proposed two different approaches for feature vector generation namely sector mean of only sal components and only cal components value of all the vectors in each sector with augmentation of extra highest-sal, zero-cal components and without augmentation of extra highest-sal, zero-cal components with sum of absolute difference and Euclidean distance [7-9] [11-14] as similarity measures. Performances of all these approaches are compared using both similarity measures.

A. Four Walsh Transform Sectors:

To get the angle in the range of 0-360 degrees, the steps as given in Table 1 are followed to separate these points into four quadrants of the complex plane. The Walsh transform of the color image is calculated in all three R, G and B planes. The complex rows representing sal components of the image and the real rows representing cal components are checked for positive and negative signs. The sal and cal Walsh values are assigned to each quadrant. as follows:

TABLE I. FOUR WALSH SECTOR FORMATION

Sign of Sal	Sign of Cal	Quadrant Assigned
+	+	I (0 – 90 Degrees)
+	-	II (90 – 180 Degrees)
-	-	III(180- 270 Degrees)
-	+	IV(270–360 Degrees)

The equation (1) is used to generate individual components to generate the feature vector of dimension 12 considering three R, G and B Planes in the sal and cal density distribution approach. However, it is observed that the density variation in 4 quadrants is very small for all the images. Thus the feature vectors have poor discretionary power and hence higher number of sectors such as 8, 12 and 16 were tried. In the case of second approach of feature vector generation i.e. individual sector mean has better discretionary power in all sectors. Sum of absolute difference measure is used to check the closeness of the query image from the database image and precision and recall are calculated to measure the overall performance of the algorithm.

B. Eight Walsh Transform Sectors:

Each quadrants formed in the previous obtained 4 sectors are individually divided into 2 sectors each considering the angle of 45 degree. In total we form 8 sectors for R,G and B planes separately as shown in the Table 2. The percentage

density distribution of sal and cal in all 8 sectors are determined using equation (1) to generate the feature vector.

TABLE 2. EIGHT WALSH SECTOR FORMATION

Quadrant of 4 Walsh sectors	Condition	New sectors Formed
I (0 – 90°)	Cal >= Sal	I (0-45 Degrees)
	Sal > Cal	II (45-90 Degrees)
II (90 – 180°)	Sal > Cal	III(90-135 Degrees)
	Cal >= Sal	IV(135-180 Degrees)
III (180- 270°)	Cal >= Sal	V (180-225 Degrees)
	Sal > Cal	VI (225-270 Degrees)
IV (270 – 360°)	Sal > Cal	VII (270-315 Degrees)
	Cal >= Sal	VIII (315-360 Degrees)

C. Twelve Walsh Transform Sectors:

Each quadrants formed in the previous section of 4 sectors are individually divided into 3 sectors each considering the angle of 30 degree. In total we form 12 sectors for R,G and B planes separately as shown in the Table 3. The percentage density distribution of sal and cal in all 12 sectors are determined using equation (1) to generate the feature vector

TABLE 3. TWELVE WALSH SECTOR FORMATION

4 Quadrants	Condition	New sectors
I (0 – 90°)	Cal >= $\sqrt{3} * \text{Sal}$	I (0-30°)
	$1/\sqrt{3} \text{ cal} \leq \text{sal} \leq \sqrt{3} \text{ cal}$	II (30-60°)
	Otherwise	III (60-90°)
II (90 – 180°)	Cal >= $\sqrt{3} * \text{Sal}$	IV (90-120°)
	$1/\sqrt{3} \text{cal} \leq \text{sal} \leq \sqrt{3} \text{cal} $	V (120-150°)
	Otherwise	VI (150-180°)
III(180-270°)	Cal >= $\sqrt{3} * \text{Sal} $	VII (180-210°)
	$1/\sqrt{3} \text{ cal} \leq \text{sal} \leq \sqrt{3} \text{cal} $	VIII(210-240°)
	Otherwise	IX (240-270°)
IV (270 – 360°)	Cal >= $\sqrt{3} * \text{Sal} $	X (270-300°)
	$1/\sqrt{3} \text{cal} \leq \text{sal} \leq \sqrt{3} \text{cal} $	XI (300-330°)
	Otherwise	XII (330-360°)

IV. RESULTS AND DISCUSSION

The sample Images of the database of 1055 images of 12 different classes such as Flower, Sunset, Barbie, Tribal, Puppy, Cartoon, Elephant, Dinosaur, Bus, Parrots, Scenery, Beach are shown in the Figure 1. The algorithm is tested by taking 5 query images from each class and then averaging the performance in terms of precision and recall over all the classes.

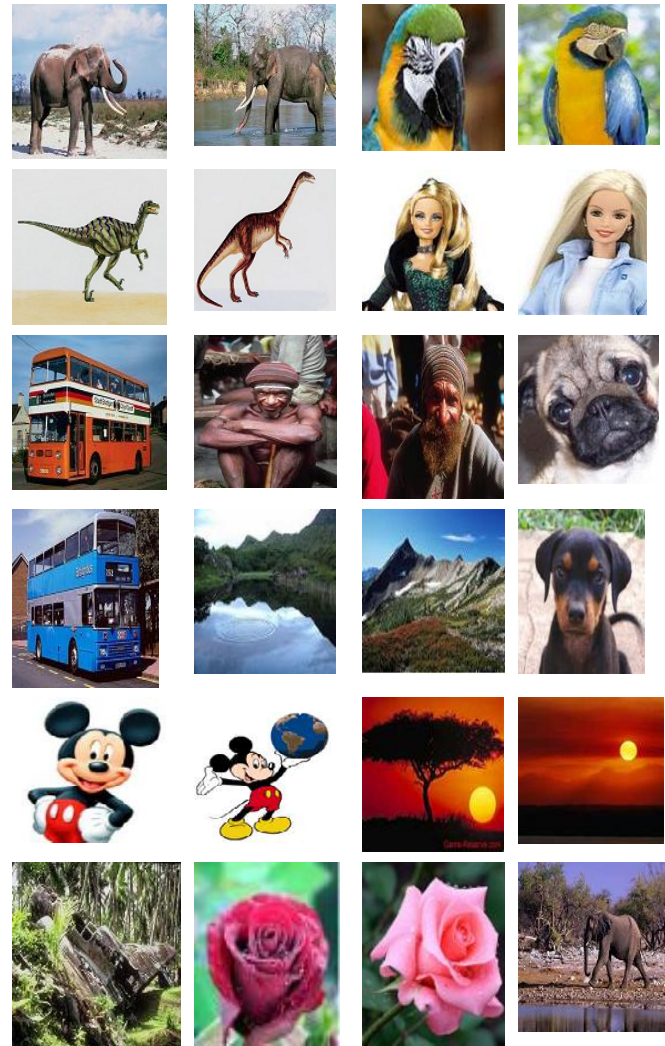


Figure 1. Sample Image Database



Figure 2. Query Image

The dinosaur class image is taken as sample query image as shown in the Figure 2. The first 21 images retrieved in the case of sector mean in 12 Walsh sector used for feature vectors and Absolute difference as similarity measure is shown in the Figure 3. It is seen that all images retrieved among first 21 images are of same class of query image i.e. dinosaur.

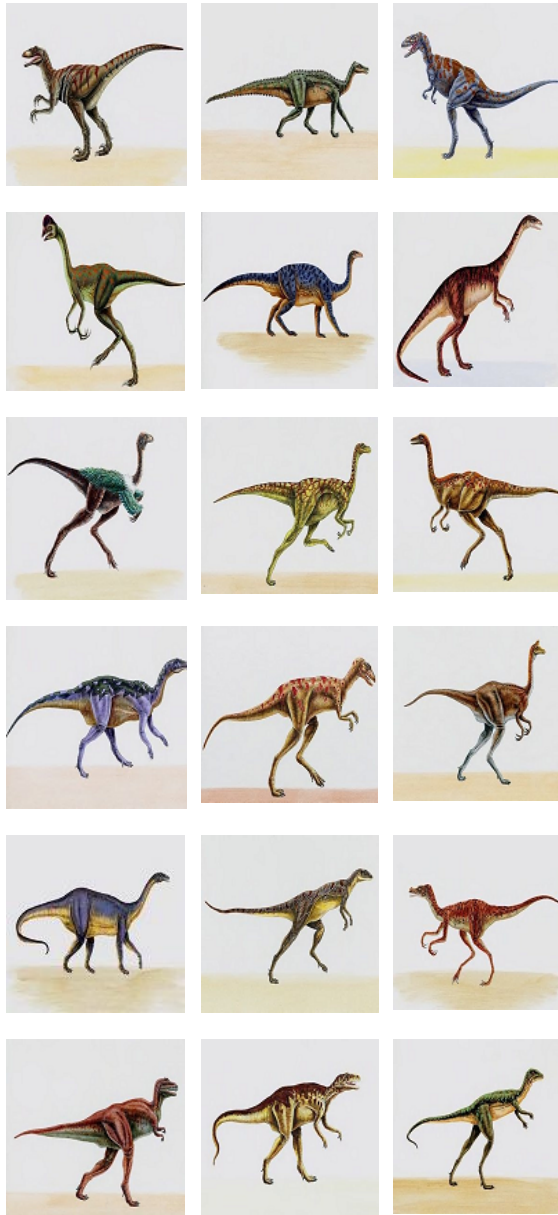


Figure 3: First 21 Retrieved Images based on individual sector mean with augmentation of zero-cal and highest-sal components of 12 Walsh Sectors with Absolute Difference as similarity measures for the query image shown in the Figure 2.

Once the feature vector is generated for all images in the database a feature database is created. A query image of each class is produced to search the database. The image with exact match gives minimum sum of absolute difference. To check the effectiveness of the work and its performance with respect to retrieval of the images we have calculated the precision and recall as given in Equations (1) and (2) below:

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total Number of images retrieved}} \quad (1)$$

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in database}} \quad (2)$$

The Figure 4 – Figure 7 shows the Overall Average Precision and Recall performance of mean of only sal components of each sectors in 4, 8, 12 and 16 Walsh Transform sectors with absolute Difference respectively. Figure 8 – Figure 11 shows the overall average cross over performance of individual sector mean of only cal components in 4, 8, 12 and 16 Walsh sectors. The comparison bar chart of cross over points of overall average of precision and recall for 4, 8, 12 and 16 sectors of with augmentation of extra zero-cal and highest-sal components with individual sector mean w.r.t. two different similarity measures namely Euclidean distance and Absolute difference is shown in the Figure 12 and Figure13. It is observed that performance of 12 sectors with extra components of zero-cal and highest-sal with consideration of only cal components of each sector is the best. The performance of absolute difference is quite closed to Euclidean distance.

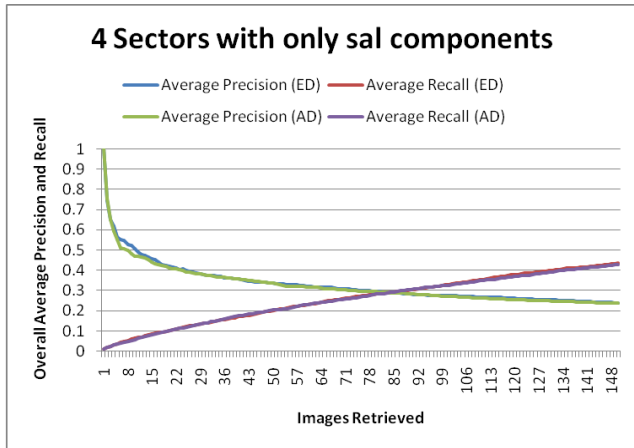


Figure 4: Overall Average Precision and Recall performance of Sector mean with only sal component in 4 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

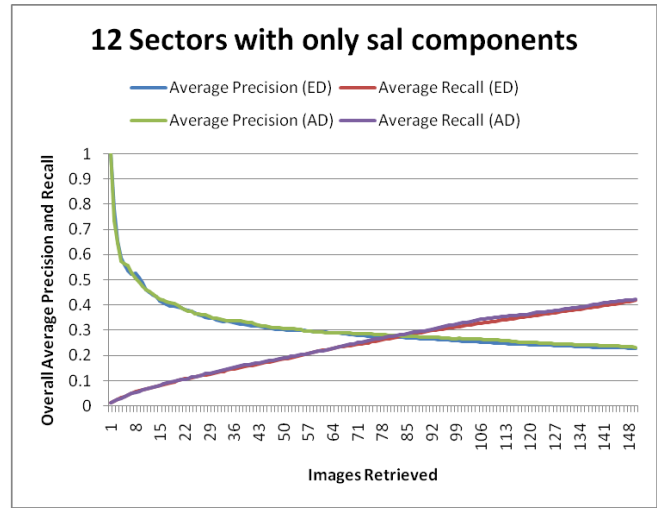


Figure 6: Overall Average Precision and Recall performance of Sector mean with only sal component in 12 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

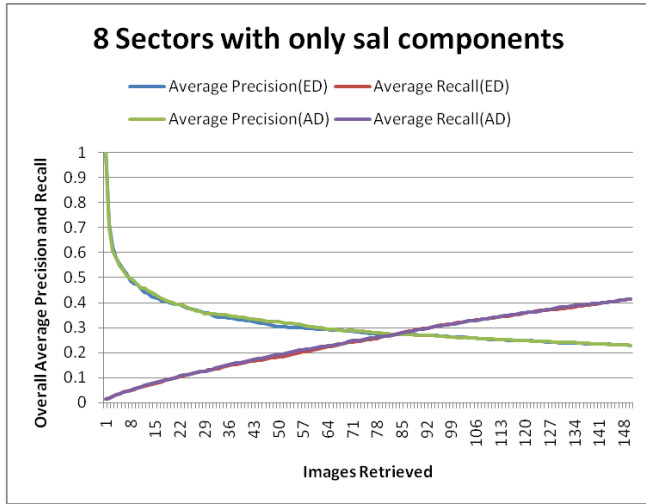


Figure 5: Overall Average Precision and Recall performance of Sector mean with only sal component in 8 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

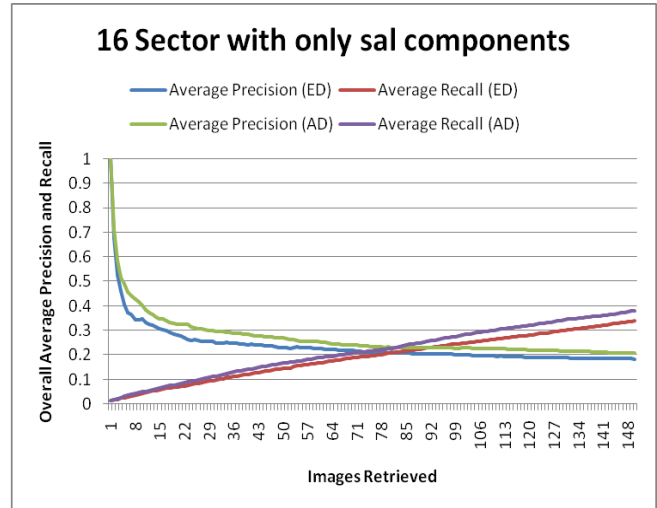


Figure 7: Overall Average Precision and Recall performance of Sector mean with only sal component in 16 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

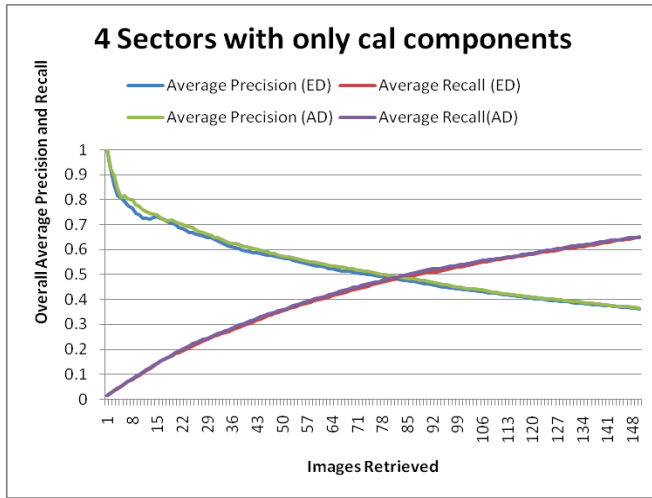


Figure 8: Overall Average Precision and Recall performance of Sector mean with only cal component in 4 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

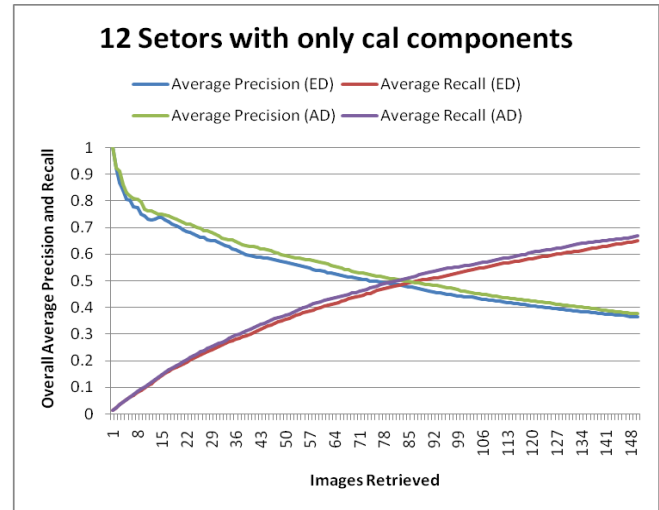


Figure 10: Overall Average Precision and Recall performance of Sector mean with only cal component in 12 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

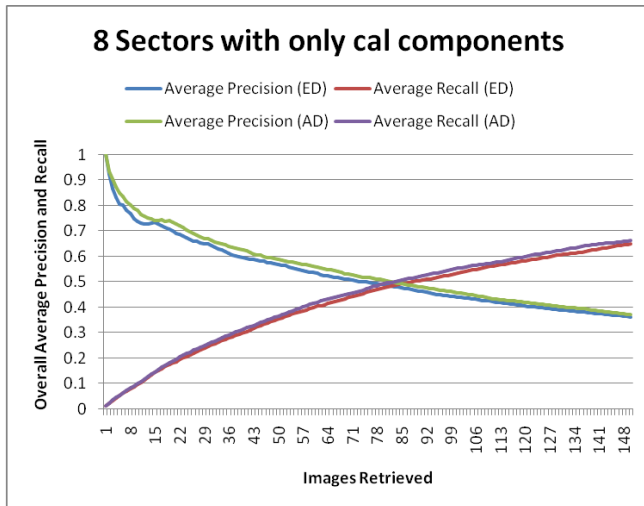


Figure 9: Overall Average Precision and Recall performance of Sector mean with only cal component in 8 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

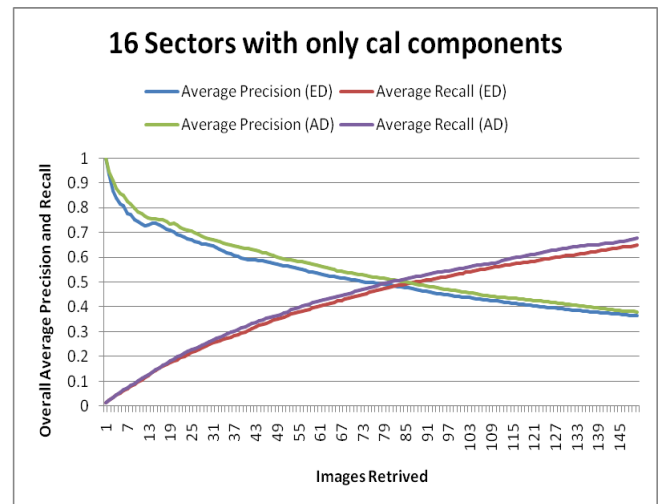


Figure 11: Overall Average Precision and Recall performance of Sector mean with only cal component in 16 Walsh Transform sectors with Absolute Difference(AD) and Euclidian Distance (ED) as similarity measures.

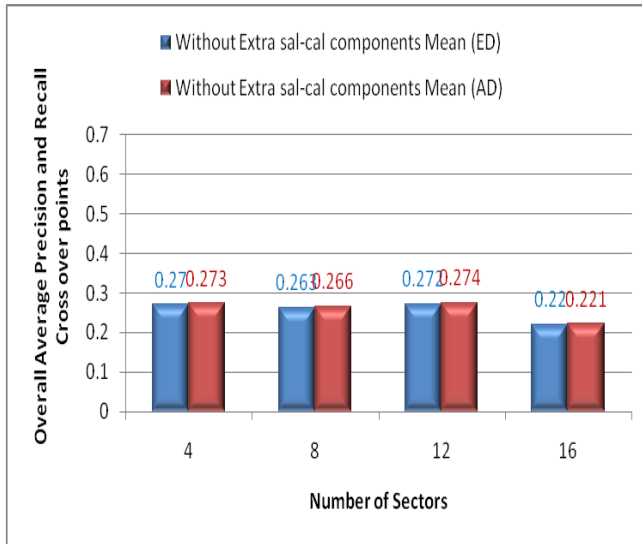


Figure 12: Comparison of Overall Precision and Recall cross over points based on individual sector mean in Walsh 4, 8, 12 and 16 sectors with Absolute Difference (AD) and Euclidean Distance (ED) as similarity measure.

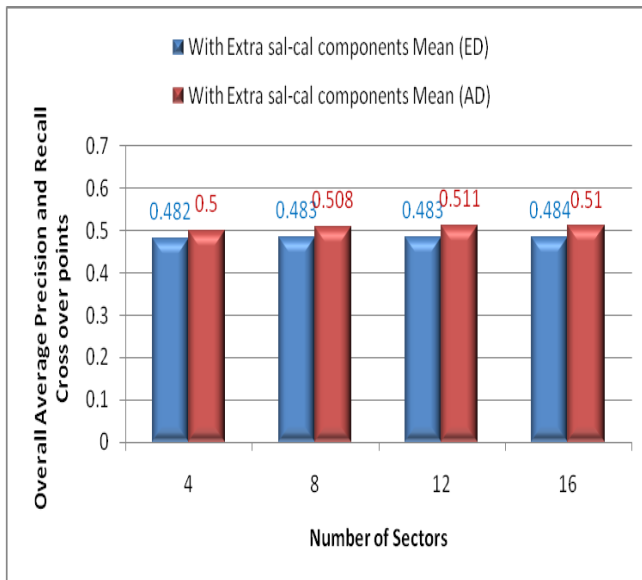


Figure 13: Comparison of Overall Precision and Recall cross over points based on individual sector mean in Walsh 4, 8, 12 and 16 sectors with Absolute Difference (AD) and Euclidean Distance (ED) as similarity measure.

V. CONCLUSION

The Innovative idea of using complex Walsh transform 4, 8, 12 and 16 sectors of the images to generate the feature vectors for content based image retrieval is proposed. We have proposed two different approaches using only sal and

only cal for feature vector generation with sum of absolute difference as similarity measuring parameter. These results are compared with Euclidian distance as similarity measure. Both these approaches are experimented with and without augmentation of extra component of zero-cal and highest-sal. The cross over point performance of overall average of precision and recall for both approaches on all applicable sectors are compared. It is found that the sector mean of only cal component with augmentation of extra component of zero-cal and highest-sal always gives the best outcome of retrieval as shown in the bar chart of the figure 13. It is also observed that sum of absolute difference is found economical similarity measuring parameter. Using Walsh transform and absolute difference as similarity measuring parameter reduces the computational complexity reducing the search time and calculation of feature vector [8][9].

REFERENCES

- [1] Kato, T., "Database architecture for content based image retrieval in Image Storage and Retrieval Systems" (Jambardino A and Niblack W eds), *Proc SPIE 2185*, pp 112-123, 1992.
- [2] John Berry and David A. Stoney "The history and development of fingerprinting," in *Advances in Fingerprint Technology*, Henry C. Lee and R. E. Gaensslen, Eds., pp. 1-40. CRC Press Florida, 2nd edition, 2001.
- [3] Emma Newham, "The biometric report," SJB Services, 1995.
- [4] H. B. Kekre, Dharendra Mishra, "Digital Image Search & Retrieval using FFT Sectors" published in proceedings of National/Asia pacific conference on Information communication and technology(NCICT 10) 5TH & 6TH March 2010.SVKM'S NMIMS MUMBAI
- [5] H.B.Kekre, Dharendra Mishra, "Content Based Image Retrieval using Weighted Hamming Distance Image hash Value" published in the proceedings of international conference on contours of computing technology pp. 305-309 (Thinkquest2010) 13th & 14th March 2010.
- [6] H.B.Kekre, Dharendra Mishra, "Digital Image Search & Retrieval using FFT Sectors of Color Images" published in International Journal of Computer Science and Engineering (IJCSIS) Vol. 02, No.02, 2010, pp.368-372 ISSN 0975-3397 available online at <http://www.enggjournals.com/ijcse/doc/IJCSE10-02-02-46.pdf>
- [7] H.B.Kekre, Dharendra Mishra, "CBIR using upper six FFT Sectors of Color Images for feature vector generation" published in International Journal of

- Engineering and Technology(IJET) Vol. 02, No. 02, 2010, 49-54 ISSN 0975-4024 available online at
<http://www.enggjournals.com/ijet/doc/IJET10-02-02-06.pdf>
- [8] H.B.Kekre, Dharendra Mishra, "Four walsh transform sectors feature vectors for image retrieval from image databases", published in international journal of computer science and information technologies (IJCSIT) Vol. 1 (2) 2010, 33-37 ISSN 0975-9646 available online at
<http://www.ijcsit.com/docs/vol1issue2/ijcsit2010010201.pdf>
- [9] H.B.Kekre, Dharendra Mishra, "Performance comparison of four, eight and twelve Walsh transform sectors feature vectors for image retrieval from image databases", published in international journal of Engineering, science and technology(IJEST) Vol.2(5) 2010, 1370-1374 ISSN 0975-5462 available online at
<http://www.ijest.info/docs/IJEST10-02-05-62.pdf>
- [10] H.B.Kekre, Dharendra Mishra, "Density distribution in Walsh Transform sectors as feature vectors for image retrieval", International journal of computer application (IJCA), ISSN NO. 975-8887, Vol.4, No.6, July 2010, pp-30-36.available online at
<http://www.ijcsonline.org/archives/volume4/number6/829-1072>.
- [11] Arun Ross, Anil Jain, James Reisman, "A hybrid fingerprint matcher," *Int'l conference on Pattern Recognition (ICPR)*, Aug 2002.
- [12] A. M. Bazen, G. T. B.Verwaaijen, S. H. Gerez, L. P. J. Veelenturf, and B. J. van der Zwaag, "A correlation-based fingerprint verification system," *Proceedings of the ProRISC2000 Workshop on Circuits, Systems and Signal Processing*, Veldhoven, Netherlands, Nov 2000.
- [13] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, "Image Retrieval using Color-Texture Features from DCT on VQ Codevectors obtained by Kekre's Fast Codebook Generation", ICGST International Journal on Graphics, Vision and Image Processing (GVIP), Available online at
<http://www.icgst.com/gvip>
- [14] H.B.Kekre, Sudeep D. Thepade, "Using YUV Color Space to Hoist the Performance of Block Truncation Coding for Image Retrieval", IEEE International Advanced Computing Conference 2009 (IACC'09), Thapar University, Patiala, INDIA, 6-7 March 2009.
- [15] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Augmented Block Truncation Coding Techniques", ACM International Conference on Advances in Computing, Communication and Control (ICAC3-2009), pp.: 384-390, 23-24 Jan 2009, Fr. Conceicao Rodrigous College of Engg., Mumbai. Available online at ACM portal.
- [16] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, "DCT Applied to Column mean and Row Mean Vectors of Image for Fingerprint Identification", International Conference on Computer Networks and Security, ICCNS-2008, 27-28 Sept 2008, Vishwakarma Institute of Technology, Pune.
- [17] H.B.Kekre, Sudeep Thepade, Archana Athawale, Anant Shah, Prathmesh Velekar, Suraj Shirke, "Walsh transform over row mean column mean using image fragmentation and energy compaction for image retrieval", International journal of computer science and engineering (IJCSE), Vol.2.No.1,S2010,47-54.
- [18] H.B.Kekre, Vinayak Bharadi, "Walsh Coefficients of the Horizontal & Vertical Pixel Distribution of Signature Template", In Proc. of Int. Conference ICIP-07, Bangalore University, Bangalore. 10-12 Aug 2007.

AUTHORS PROFILE



Dr. H. B. Kekre has received B.E. (Hons.) in Telecomm. Engg. from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg. (Electrical Engg.) from University of Ottawa in 1965 and Ph.D.(System Identification) from IIT Bombay in 1970. He has worked Over 35 years as Faculty and H.O.D. Computer science and Engg. At IIT Bombay. From last 13 years working as a professor in Dept. of Computer Engg. at Thadomal Shahani Engg. College, Mumbai. He is currently Senior Professor working with Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS University vile parle west Mumbai. He has guided 17 PhD.s 150 M.E./M.Tech Projects and several B.E./B.Tech Projects. His areas of interest are Digital signal processing, Image Processing and computer networking. He has more than 300 papers in National/International Conferences/Journals to his credit. Recently ten students working under his guidance have received the best paper awards. Currently he is guiding 8 Ph.D. Students. Two of his Students have recently completed Ph. D. He is life member of ISTE and Fellow of IETE.



Dhirendra S.Mishra has received his BE (Computer Engg) degree from University of Mumbai in 2002. Completed his M.E. (Computer Engg) from Thadomal shahani Engg. College, Mumbai, University of Mumbai. He is PhD Research Scholar and working

as Assistant Professor in Computer Engineering department of Mukesh Patel School of Technology Management and Engineering, SVKM's NMIMS University, Mumbai, INDIA. He is life member of Indian Society of Technical education (ISTE), Member of International association of computer science and information technology (IACSIT), Singapore, Member of International association of Engineers (IAENG). His areas of interests are Image Processing, Operating systems, Information Storage and Management

Supervised Learning Approach for Predicting the Presence of Seizure in Human Brain

Sivagami P, Sujitha V
M.Phil Research Scholar
PSGR Krishnammal College for Women
Coimbatore, India

Vijaya MS
Associate Professor and Head
GRG School of Applied Computer Technology
PSGR Krishnammal College for Women
Coimbatore, India.

Abstract— Seizure is a synchronous neuronal activity in the brain. It is a physical change in behavior that occurs after an episode of abnormal electrical activity in the brain. Normally two diagnostic tests namely Electroencephalogram (EEG) and Magnetic Resonance Imaging (MRI) are used to diagnose the presence of seizure. The sensitivity of the human eye in interpreting large numbers of images decreases with increasing number of cases. Hence, it is essential to automate the accurate prediction of seizure in patients. In this paper supervised learning approaches has been employed to model the prediction task and the experiments show about 94% high prediction accuracy.

Keywords— Seizure; Support vector machine; K-NN; Naïve Bayes; J48

I. INTRODUCTION

Seizure is defined as a transient symptom of "abnormal excessive in the brain". Seizures can cause involuntary changes in body movement or function, sensation, awareness, or behavior. It is an abnormal, unregulated electrical discharge that occurs within the brain's cortical grey matter and transiently interrupts normal brain function [1]. Based on the physiological characteristics of seizure and the abnormality in the brain, the kind of seizure is determined. Seizure is broadly classified into absence seizure, simple partial, complex partial and general seizure. Absence seizure is a brief episode of staring. It usually begins in childhood between ages 4 and 14. Simple partial seizure affects only a small region of the brain, often the hippocampus. Complex partial seizure usually starts in a small area of the temporal lobe or frontal lobe of the brain. General seizure affects the entire brain.

Various diagnostic techniques normally employed for patients are Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and PET (Positron Emission Tomography). Magnetic Resonance Imaging (MRI) is used as a valuable tool and widely used in the clinical and surgical environment for seizure identification because of its characteristics like superior soft tissue differentiation, high spatial resolution and contrast. Magnetic Resonance Images are examined by radiologists based on visual interpretation of the films to identify the presence of seizure.

Machine learning is a technique which can discover previously unknown regularities and trends in diverse datasets [2]. Today machine learning provides several indispensable tools for intelligent data analysis. Machine learning technology is currently well suited for analyzing medical data and empirical results reveal that the machine learning systems are highly efficient and could significantly reduce the computational complexities.

Yong Fan developed a method for diagnosis of brain abnormality using both structural and functional MRI images [3]. Christian E. Elger, Klaus Lehnertz developed a seizure prediction by non-linear time series analysis of brain electrical activity [4]. J.W.Wheless, L.J.Willmore, J.I.Breier, M.Kataki, J.R.Smith, D.W.King provides the comparison of Magnetoencephalography, MRI, and V-EEG in Patients Evaluated for Epilepsy Surgery [5]. William D.S. Killgorea, Guila Glossera, Daniel J. Casasantoa, Jacqueline A. Frencha, David C. Alsoph, John A. Detreab provide a complementary information for predicting post-operative seizure control [6].

The motivation behind the research reported in this paper is to predict the presence of seizure in human brain. Machine learning techniques are employed here to model the seizure prediction problem as classification task to facilitate physician for accurate prediction of seizure presence. In this paper supervised learning algorithms are made use of for the automated prediction of type of seizure.

II. PROPOSED METHODOLOGY

The proposed methodology models the seizure prediction as a classification task and provides a convenient solution by using supervised classification algorithms. Descriptive features of MRI image such as energy, entropy, mean, standard deviation, contrast, homogeneity of grey scale image have been extracted and used for training. The model is trained using training datasets and the trained model is built. Finally the trained model is used to predict the type of seizure.

The proposed model is shown in Figure.1.

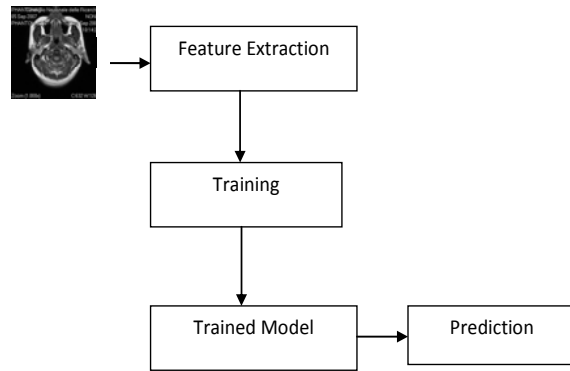


Figure 1. The Proposed model

A. Image Acquisition

A magnetic resonance imaging (MRI) scan of the patient's brain is a noninvasive method to create detailed pictures of the brain and surrounding nerve tissues. MRI uses powerful magnets and radio waves. The MRI scanner contains the magnet. The magnetic field produced by an MRI is about 10 thousand times greater than the earth's. The magnetic field forces hydrogen atoms in the body to line up in a certain way. When radio waves are sent toward the lined-up hydrogen atoms, it bounces back and a computer records the signal. Different types of tissues send back different signals.

The MRI dataset consisting of MRI scans images of 350 patients of five types namely Normal, Absence Seizure, Simple Partial Seizure, Complex Partial Seizure and General Seizure are taken into consideration.

B. Feature Extraction

The purpose of feature extraction is to reduce the original data set by measuring certain properties or features that distinguish one input pattern from another. A brain MRI slices is given as an input. The various features based on statistical, grey level co-occurrence matrix and grey level run-length matrix from the MRI is extracted. The extracted features provide the characteristics of the input type to the classifier by considering the description of the relevant properties of the image into a feature space.

The statistical features based on image intensity are mean variance, skewness and kurtosis. The grey level co-occurrence matrices (GLCM) features such as Contrast, Homogeneity, Correlation, Energy, Entropy and the features of grey level run length matrices (GLRLM) such as Short run emphasis, Long run emphasis, Grey level distribution, Run-length distribution, Run percentage, Low grey level run emphasis, High grey level run emphasis are used to investigate the adequacy for the discrimination of the presence of seizure. Table I shows the features of MRI of a human brain.

TABLE I. FEATURES OF MRI

Statistical	Grey Level Co-occurrence Matrix	Grey Level Run Length Matrix
Mean Variance Skewness Kurtosis	Contrast Homogeneity Correlation Energy Entropy	Short run emphasis Long run emphasis Grey level distribution Run length distribution Run percentage Low grey level run emphasis High grey level run emphasis

1) Grey Level Co-occurrence Matrix(GLCM)

The GLCM is defined as a tabulation of different combinations of pixel brightness values (grey levels) occur in an image. The texture filter functions provide a statistical view of texture based on the image histogram. This function provides useful information about the texture of an image but does not provide information about shape, i.e., the spatial relationships of pixels in an image.

The features corresponding to GLCM statistics and their description are:

- Contrast - Measures the local variations in the grey-level co-occurrence matrix.
- Homogeneity - Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.
- Correlation - Measures the joint probability occurrence of the specified pixel pairs.
- Energy - Provides the sum of squared elements in the GLCM. Also known as uniformity or the angular second moment.
- Entropy - statistical measure of randomness.

2) Grey Level Run Length Matrix(GLRLM)

The GLRLM is based on computing the number of grey-level runs of various lengths. A grey-level run is a set of consecutive and collinear pixel points having the same grey level value. The length of the run is the number of pixel points in the run [7]. Seven features are extracted from this matrix.

C. Supervised Classification Algorithms

Supervised learning is a machine learning technique for deducing a function from training data. The training data consist of pairs of input objects and desired outputs. The output of the function can predict a class label of the input object called classification. The task of the supervised learner is to predict the value of the function for any valid input object after having seen a number of training examples i.e. pairs of input and target output. The supervised classification techniques namely, support vector machine, decision tree

induction, Naive Bayes and k-nn are employed in seizure prediction modeling.

1) Support Vector Machine

The machine is presented with a set of training examples, (x_i, y_i) where the x_i is the real world data instances and the y_i are the labels indicating which class the instance belongs to. For the two class pattern recognition problem, $y_i = +1$ or $y_i = -1$. A training example (x_i, y_i) is called positive if $y_i = +1$ and negative otherwise [6]. SVMs construct a hyper plane that separates two classes and tries to achieve maximum separation between the classes. Separating the classes with a large margin minimizes a bound on the expected generalization error.

The simplest model of SVM called Maximal Margin classifier, constructs a linear separator (an optimal hyper plane) given by $w^T x - y = 0$ between two classes of examples. The free parameters are a vector of weights w which is orthogonal to the hyper plane and a threshold value. These parameters are obtained by solving the following optimization problem using Lagrangian duality.

$$\text{Minimize} = \frac{1}{2} \|w\|^2$$

$$\text{subject to } D_{ii} (w^T x_i - \gamma) \geq 1, i = 1, \dots, l. \quad (1)$$

where D_{ii} corresponds to class labels $+1$ and -1 . The instances with non null weights are called support vectors. In the presence of outliers and wrongly classified training examples it may be useful to allow some training errors in order to avoid over fitting. A vector of slack variables ξ_i that measure the amount of violation of the constraints is introduced and the optimization problem referred to as soft margin is given below. In this formulation the contribution to the objective function of margin maximization and training errors can be balanced through the use of regularization parameter C . The following decision rule is used to correctly predict the class of new instance with a minimum error.

$$f(x) = \text{sgn}[w^T x - \gamma] \quad (2)$$

The advantage of the dual formulation is that it permits an efficient learning of non-linear SVM separators, by introducing kernel functions. Technically, a kernel function calculates a dot product between two vectors that have been (non-linearly) mapped into a high dimensional feature space [8]. Since there is no need to perform this mapping explicitly, the training is still feasible although the dimension of the real feature space can be very high or even infinite. The parameters are obtained by solving the following non linear SVM formulation (in matrix form),

$$\text{Minimize LD}(u) = \frac{1}{2} u^T Q u - e^T u$$

$$d^T u = 0 \quad 0 \leq u \leq C e \quad (3)$$

where K - the Kernel Matrix. $Q = DKD$.

The Kernel function K (AAT) (polynomial or Gaussian) is used to construct hyperplane in the feature space, which separates two classes linearly, by performing computations in the input space.

$$f(x) = \text{sgn}(K(x, x_i^T) * u - \gamma) \quad (4)$$

where u - the Lagrangian multipliers. In general larger the margins will lower the generalization error of the classifier.

2) Naïve Bayes

Naïve Bayes is one of the simplest probabilistic classifiers. The model constructed by this algorithm is a set of probabilities. Each member of this set corresponds to the probability that a specific feature f_i appear in the instances of class c , i.e., $P(f_i | c)$. These probabilities are estimated by counting the frequency of each feature value in the instances of a class in the training set. Given a new instance, the classifier estimates the probability that the instance belongs to a specific class, based on the product of the individual conditional probabilities for the feature values in the instance. The exact calculation uses bayes theorem and this is the reason why the algorithm is called a bayes classifier.

3) K-NN

K -nearest neighbor algorithms are only slightly more complex. The k nearest neighbor of the new instance is retrieved and whichever class is predominant amongst them is given as the new instance's classification. K -nearest neighbor is a supervised learning algorithm where the result of new instance query is classified based on majority of K -nearest neighbor category [9]. The purpose of this algorithm is to classify a new object based on attributes and training samples. The classifiers do not use any model to fit and only based on memory.

4) J48 Decision Tree Induction

J48 algorithm is an implementation of the C4.5 decision tree learner. This implementation produces decision tree models. The algorithm uses the greedy technique to induce decision trees for classification [10]. A decision-tree model is built by analyzing training data and the model is used to classify unseen data. J48 generates decision trees, the nodes of which evaluate the existence or significance of individual features.

III. EXPERIMENTAL SETUP

The seizure data analysis and Prediction has been carried out using WEKA and SVM^{light} for machine learning.

WEKA is a collection of machine learning algorithms for data mining tasks [11]. SVM^{light} provides the extensive support for the whole process of experiment including preparing the input data, evaluating learning schemes statistically and visualizing the input data and the result of learning.

The dataset is trained using SVM with most commonly used kernels linear, polynomial and RBF, with different parameter settings for d , γ and C –regularization parameter. The parameters d and γ are associated with polynomial kernel and RBF kernel respectively. Image processing toolbox of Matlab has been used for MRI feature extraction. The datasets are grouped into five broad classes namely Normal, Absence Seizure, Simple Partial Seizure, Complex Partial Seizure and General Seizure to facilitate their use in experimentally determining the presence of seizure in MRI. The seizure dataset has 17 attributes, there are 350 instances, and as indicated above, 5 classes. Supervised classification algorithms such as support vector machine, decision tree induction, naïve bayes and K-NN are applied for training. Support vector machine learning is implemented using SVM light. Decision tree induction, Naïve Bayes and K-NN are implemented using WEKA. The performance of the trained models has been evaluated using 10 fold cross validation and their results are compared.

IV. RESULTS

The results of the experiments are summarized in Table II. Prediction accuracy and learning time are the parameters considered for performance evaluation. Prediction accuracy is the ratio of number of correctly classified instances and the total number of instances. Learning time is the time taken to build the model on the dataset.

A. Classification using SVM

The performance of the three kinds of SVMs with linear, polynomial and RBF kernels are evaluated based on the prediction accuracy and the results are shown in Table II.

TABLE II. SVM - LINEAR, POLYNOMIAL, RBF KERNELS

SVM Kernels	C=1		C=2		C=3		C=4	
Linear	74		76		72		79	
Polynomial (d)	1	2	1	2	1	2	1	2
	79	81.2	82	80	86	84	74	75
RBF (g)	0.5	1	0.5	1	0.5	1	0.5	1
	92	94	93	92	95	97	94	95

Table III shows the average performance of the SVM based classification model in terms of predictive accuracy and

depicted the same in Figure.2.

TABLE III. AVERAGE PERFORMANCE OF THREE MODELS

Kernel Type	Prediction Accuracy(%)
Linear	75
Polynomial	80
RBF	94

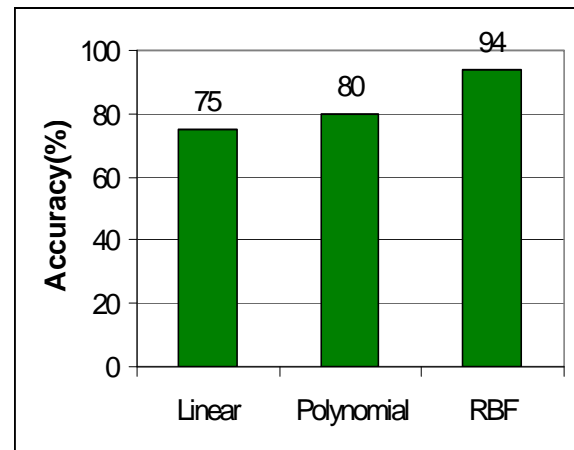


Figure 2. Comparing Prediction Accuracy of SVM Kernels

The predictive accuracy shown by SVM with RBF kernel with parameter $C=3$ and $\gamma=2$ is higher than the linear and polynomial kernel.

B. Classification using WEKA

The results of the experiments are summarized in Table IV and V.

TABLE IV. PREDICTIVE PERFORMANCE

Classifiers	Evaluation Criteria			
	Learning Time (secs)	Correctly classified instances	Incorrectly classified instances	Prediction accuracy (%)
Naïve Bayes	0.03	272	68	80
K-NN	0.02	276	64	81.17
J48	0.09	293	47	86.17

V. CONCLUSION

This paper describes the modeling of the seizure prediction task as classification and the implementation of trained model using supervised learning techniques namely, Support vector machine, Decision tree induction, Naive Bayes and K-NN. The performance of the trained models are evaluated using 10 fold cross validation based on prediction accuracy and learning time and the results are compared. It is observed that about 94% high predictive accuracy is shown by the seizure prediction model. As far as the seizure prediction is concerned, the predictive accuracy plays major role in determining the performance of the model than the learning time. The comparative results indicate that support vector machine yield a better performance when compared to other supervised classification algorithms. Due to wide variability in the dataset, machine learning techniques are effective than the statistical approach in improving the predictive accuracy.

TABLE V. COMPARISON OF ESTIMATES

Evaluation criteria	Classifiers		
Kappa Statistic	Naïve Bayes	K-NN	J48
Mean Absolute Error	0.7468	0.7614	0.8235
Root Mean Squared Error	0.2716	0.266	0.2284
Relative Absolute Error	26.1099	24.2978	21.2592
Root Relative Squared Error	68.428	67.0142	57.549

The performances of the three models are illustrated in Figure 3 and 4.

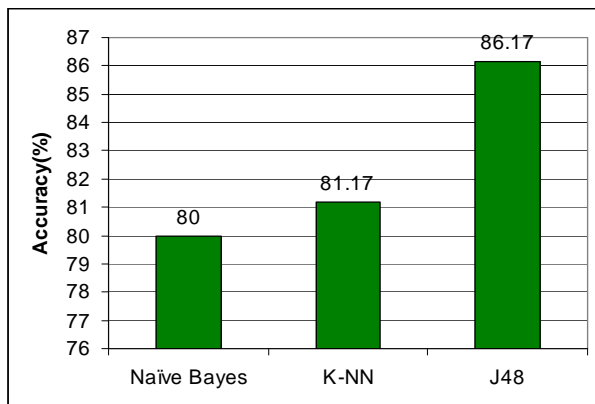


Figure 3. Comparing Prediction Accuracy

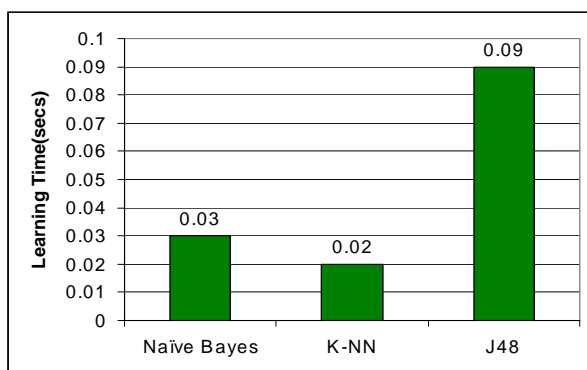


Figure 4. Comparing Learning Time

The time taken to build the model and the prediction accuracy is high in J48 when compared to other two algorithms in WEKA environment.

ACKNOWLEDGMENT

The authors would like to thank the Management and Acura Scan Centre, Coimbatore for providing the MRI data.

REFERENCES

- [1] Robin cook, "Seizure" Berkley Pub Group, 2004.
- [2] Karpagavalli S, Jamuna KS, and Vijaya MS, "Machine Learning Approach for pre operative anaesthetic risk Prediction", International Journal of Recent Trends in Engineering, Vol. 1. No.2, May 2009.
- [3] Yong fan, "Multivariate examination of brain abnormality using both structural and functional MRI", Neuroimaging, elsevier, vol 36 issue 4 pp 1189-1199, 2007
- [4] Christian E. Elger, Klaus Lehnertz, "Seizure prediction by non-linear time series analysis of brain electrical activity" European Journal of Neuroscience Vol 10, Issue 2, pages 786-789, February 1998.
- [5] J. W. Wheless, L. J. Willmore, J. I. Breier, M. Katakai, J. R. Smith, D. W. King, "A Comparison of Magnetoencephalography, MRI, and V-EEG in Patients Evaluated for Epilepsy Surgery", Epilepsia, Vol 40, Issue 7, pages 931-941, July 1999.
- [6] William D.S. Killgorea, Guila Glossera, Daniel J. Casasantoa, Jacqueline A. Frencha, David C. Alsopb, John A. Detreab, Functional MRI and the Wada test provide complementary information for predicting post-operative seizure control, Seizure, pp 450-455, Dec 1999
- [7] Galloway M. "Texture analysis using grey level runs lengths", Comp Graph Image Process, pp.72-179, 1975.
- [8] Nello Cristianini and John Shawe-Taylor. "An Introduction to Support Vector Machines and other kernel-based learning methods" Cambridge University Press, 2000.
- [9] Teknomo, Kardi. K-Nearest Neighbors Tutorial
- [10] M. Chen, A. X. Zheng, J. Lloyd, M. I. Jordan, and E. Brewer, "Failure diagnosis using decision trees", In Proc. IEEE ICAC, 2004.
- [11] Ian H. Witten, Eibe Frank, Len Trigg, Mark Hall, Geoffrey Holmes, Sally Jo Cunningham, "Weka: Practical Machine Learning Tools and Techniques with Java Implementations", 1999.

Approximate String Search for Bangla: Phonetic and Semantic Standpoint

Adeeb Ahmed

Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

Abdullah Al Helal

Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

Abstract— Despite the improvement in the field of approximate string search, insignificant research was performed for Bangla string matching. Approximate string search has a great deal of interest in spellchecking, query relaxation or interactive search. In our work, we proposed a method for Bangla string search which is specially modified considering Bangla spelling rules and grammar. Rather than simple string matching, special emphasis was given to make sure that words possessing relevant meaning are not ignored due to its inflected form. Moreover, phonetic matching was also emphasized for the purpose.

Keywords- Approximate string search; Bangla search; Levenshtein distance; query relaxation; spelling suggestion; case ending

I. INTRODUCTION

Consider a collection of strings named 'Database' and a query string named 'Queryword'. We need to find all the substrings in 'Database' which possess 'similarity' with the query string 'Queryword', and sort them according to their similarity with 'Queryword'. Now, the real challenge is to define the term 'similarity'. Different methods have been proposed for this purpose [1]–[6]. Different functions were used for finding the similarity between strings such as Levenshtein distance [7], cosine similarity [5] or Jaccard coefficient [8].

But these distances alone are not capable of dealing with common spelling mistakes made by human. Especially in Bangla, words may lose their original form when used inside a sentence as Bangla is a highly inflected language.. Considering these alterations is far beyond the scope of these functions alone. In this work, we have taken two different matters into account for the approximate search. First, the common spelling mistakes made by human in Bangla. For this purpose Bangla phonetic was studied and any mismatch between similar sounding letters was ignored. Being an extremely rich language, Bangla possess more than one characters for various similar sounding voiced and unvoiced sounds. From phonetic standpoint, they could easily been represented by a single character. Due to almost similar auditory sensation, these similar sounding letters often creates confusion and causes spelling mistakes. Moreover, in some cases, different spellings of a single word are accepted. In finding the approximate

string search, these common spelling errors must be studied carefully.

The second factor that must be taken into account is more important for implementation of query relaxation. Due to the grammatical rules in Bangla, most often words lose their original form inside a sentence due to different forms of inflections. These inflections may be classified into various groups like tense ending, case ending, personal ending, imperative ending, etc [9]. Among the various forms of inflections, case ending is responsible for alteration of nouns and pronouns. Noticeable facts about these types of inflections are that, they are unavoidable in sentence formation and they cause insignificant changes in the meaning of the words. As proper nouns and nouns together constitutes over 70% of the query terms on web [10], considering the word inflection due to case ending is extremely important for Bangla search.

II. PRELIMINARIES

Among various functions for computing the similarity between two different strings, Levenshtein distance or edit distance is an accepted one. Levenshtein distance between two strings is defined as the minimum number of operations (substitution, insertion or deletion) required for converting from one string to another. Now let us look carefully about the performance of edit distance in Bangla string matching. Here we assume the Bangla text is encoded using Unicode [11]. As stated earlier, Bangla contains several similar sounding characters which often introduce confusion; we study the spelling of the word 'BANGLA' itself. The word 'BANGLA' can be spelled in two different ways 'বাংলা' and 'বাঙলা'. If we simply consider the Levenshtein distance or edit distance, we get the value '1' (substitute ং with ঙ). Now let us consider the edit distance between the words 'বাংলা' and 'বালা' which means 'BANGLE' in English, a completely different word. A simple insertion of ং transforms 'বালা' into 'বাংলা', resulting an equal edit distance compared with the previous pair. But from phonetic point of view, 'বাংলা' is much closer to 'বাঙলা'. And on this case possess the same meaning. So before computing the conventional edit distance, these factors must be considered. Understanding the second factor, the inflections of words due to case ending requires slight knowledge on Bangla

grammar. To make the things easier, here is a brief description about the alteration process. Like prepositions used in English; a consonant (র, য় etc) , a dependent vowel (ে) [11] or both (রে, তে) may be added at the end of a word in Bangla [12]. The most troubling part is, unlike English, these additional dependent vowels or consonants merge with the words making itself an integral part of the word. As these inflections do not make any significant change in the meaning of the word, rather they are used to embed the word inside a sentence, these inflections should be considered carefully in case of a query search used for web. A simple example may help to clarify the necessity of considering the case ending. Consider someone is willing to know about the capital of Bangladesh, that is DHAKA. In Bangla it is spelled 'ঢাকা'. In Table I, some words are stated with their respective meanings and edit distances with the word of interest.

TABLE I. PERFORMANCE OF LEVENSHTIN DISTANCE FOR BANGLA

Serial	Word	English Meaning	Edit distance
1	ঢাকা	Dhaka, Capital of Bangladesh	0
2	ঢাক	Drum (Musical instrument)	1
3	ঢাকী	Drummer	1
4	টাকা	Currency of Bangladesh	1
5	ডাকা	To call	1
6	পাকা	Ripe	1
7	ঢাকায়	In Dhaka	1
8	ঢাকার	Of Dhaka	1
9	ঢাকাকে	To Dhaka (used for addressing)	2
10	ঢাকাতে	In Dhaka	2

In the list, words from 2 to 8, all having the same edit distance, would be treated with equal importance as being similar to the word of interest. But from semantic standpoint, the words numbered from 7 to 10 having the true information about the capital of Bangladesh, should be given preference. For words having puzzling spelling rules, simultaneous occurrence of an inflection due to case ending and a spelling mistake may lead to a higher edit distance, resulting in an undesirable outcome. These facts motivate us to perform an additional task before calculating the conventional edit distance for Bangla approximate search.

III. METHOD

For any string searching algorithm, one of the most important factors is running time, especially for those applications adopting a web based service model. We are assuming to have a large list of words to perform the search operation on which is evident both for a dictionary search or web query. So, we propose a method which consists of two major stages.

- Fast filtering
- Computation of modified edit distance

A. Fast Filtering

Since we are considering to have a large amount of data to search on, the best idea would be to quickly discard large part of the text using a computationally efficient method. This is called filtering and various methods like n-gram[13] or spaced seeds[14] have been proposed for serving the purpose. In our work, we applied a fast filtering method which has similarity with n-gram method (with n=1), and has a very simple form. A filter is said to be lossless if it does not discard any potential match during its operation. To be on the safe side and avoid losing a probable matching word, we have adopted a filter which acts in a defensive manner rather than being too aggressive to discard large amount of text on the first run. In the filtering process following steps were performed.

1) *Length Matching*: If query string has N characters, then only words having length between L1 and L2 inclusive will be considered for the next step, where

$$L_1 = N - N/2$$

$$L_2 = N + N/2$$

L1, L2 rounded to the nearest integer.

On this stage, large amount of words are discarded with little computational cost. A comparatively large margin is used for the words to pass the filter. This is due to the fact that, in Bangla a word within a sentence can be augmented by case ending (e.g. 'ঢাকা' may become 'ঢাকাকে', see Table 1) and this will result in a longer word. On the other hand, words may have shorter form due to some variation in spelling (e.g. sometimes 'া', 'ে' are ignored).

2) *Coarse Distance Matching*: Only the words qualifying in the first stage are considered for this stage. In this step, a similarity between the query word and the searched word is measured by comparing the number of occurrence of different characters in the two words. For serving the purpose, a one dimensional vector of length k is used, where k is the number of possible characters (including dependent vowels) in Bangla. For the query word, the vector CQ is computed before starting the process. Say,

$$C_Q = [c_{Q1} \ c_{Q2} \ c_{Q3} \ \dots \ c_{Qk}] \quad (1)$$

Where C_{Qn} = number of occurrence of n^{th} character on the query word.

Similarly C_S is computed for the searched word.

$$C_S = [c_{S1} \ c_{S2} \ c_{S3} \ \dots \ c_{Sk}] \quad (2)$$

Now, coarse distance is computed by the equation

$$C_D = \sum_{n=1}^k |c_{Qn} - c_{Sn}| \quad (3)$$

Higher the value of C_D indicates higher difference between the two words. A threshold is set for the words to pass through the filter. In this case also, moderately large threshold is used to ensure lossless filtering. Simple assumption suggests us to consider a threshold value proportional to the length of the query string as longer string can contain more errors. Only words which result C_D values lower than the threshold value are considered for the next stage.

B. Computing Modified Edit Distance

The focal part of our work is to compute a distance between words which is aptly intelligent to distinguish human error in searching and the inflections introduced in words when used in sentences. To do the job, before computing the traditional distance, some modifications are made over the words. First stage of modification takes into account the phonetic similarities between words and exploiting the phonetic resemblances, some manipulation and simplification of spellings are done. On the second stage of modification, semantic similarities are considered and alterations are done to make sure that semantically similar words in the database produce lower distance with the query word compared to the other words.

1) *Spelling Modifications*: As stated earlier, two different features are emphasized during the approximate string search in our work. First, spelling mistake due to similar sounding letters (e.g. ‘স’, ‘শ’, ‘ষ’) should be considered. This problem also exists in English and various approximate string matching algorithm built for English used a variety of phonetic methods like Soundex [15] or PHONIX [16]. Among all the methods, Soundex is the oldest. It was particularly developed for English. Soundex replaces the 26 different letters in English by a set of 7 disjoint sets only considering their phonetic similarity. Vowels are completely ignored and not taken on the computations. PHONIX is also similar to Soundex but little modification is done prior to mapping of words. But for Bangla, mapping of word to such a small number of sets and ignoring the vowel may bring up unacceptable result. Due to the word structure, small variation in a dependent or independent vowel may produce a completely new set of words with different meanings. Refer to Table 1, there is only a difference of one dependent vowel among words 1, 2 and 3. This proves the impracticality of ignoring the vowels for implementing Soundex in Bangla. Furthermore, in Soundex, the English letters are mapped into only 7 disjoint sets which demands the mapping of hardly similar sounding letters to map into the same set (eg. D, T are treated equally both in Soundex or PHONIX). But careful observation of Bangla lexicon reveals numerous words which are comparable from phonetic standpoint (eg. word 1, 4 and 5 in Table 1). Moreover, due to implementation of fast filtering in the first stage, we expect to have relatively smaller number of words. This eliminates the need for a highly computationally efficient matching. Considering these details, we used a rather conservative conversion, only by converting the phonetically similar characters, keeping most of the words unchanged (Table II).

There is a small listing of ignored characters which are often disregarded commonly. A simple example may clarify the procedure. Consider a Bangla word ‘বিভীষণ’. After conversion it will become ‘বিভিসন’.

TABLE II. MAPPING OF WORDS FOR SPELLING MODIFICATION

Characters in the original word	Converted characters
্, ঁ, ো	Ignored
ঞ, ং	ঙ
শ, ষ	স
ণ	ন
ড, ঢ	র
য	জ
ং	ত
ঈ	ই
উ	উ
ী	ি
ু	ু
Rest of the characters	Unchanged

2) *Case Ending Consideration*: The second form of modification is particularly required for web query or database search. As explained earlier, the Bangla words undergo various inflections. Due to, greater importance of nouns and pronouns in web search, in our work we only modify the inflections applied over nouns and pronouns, that is inflections due to case ending.

To make things even complicated, most of the case ending terms in Bangla words are integrated with the original words making it even harder to deal with. But fortunately, there are limited numbers of case ending terms listed in Table III used in Bangla, and by using proper logic; these can be identified most of the time.

TABLE III. LIST OF CASE ENDINGS USED IN BANGLA

Group-1	কে, রে, এরে, এ, য়, তে, এতে, র, এর
Group-2	দ্বারা, দিয়া, কর্তৃক, হইতে, থেকে, চেয়ে

In Table III, case endings listed as Group-2 do not unite with the original words and thus not of our concern. Only Group-1 case endings would be considered. Here, a noteworthy thing is that, the case endings written in Table III are not in the exact form how they exist inside the word. For ease of reading, all the case endings are written using independent vowels (এ, এতে, etc) in Table III. But when used with words these independent vowels would be replaced by dependent vowels (e.g. এ with ে). As an example, when case ending ‘এর’ is used with a word ‘সকাল’, the word after inflection would be

সকাল + এর = সকালের

In the modification process, the elimination of the case ending may seem to result a faulty exclusion. When the word already free of case ending ends with a set of characters from Group-1, in Table III, the fault may occur. But fortunately, careful inspection of Bangla lexicon confirms existence of handful words containing exact matches of case endings at the end. And even if the case arises, (e.g. the query word is 'সরকার', containing 'র' at the end), it is very improbable that after performing the faulty exclusion (in this case the modified word would be 'সরকা') there will be any word in the wordlist being searched that would exactly match the modified query word. If the wordlist being searched for contains the word 'সরকার', same faulty exclusion will take place for that word also, which will therefore result zero edit distance for the pair.

3) *Distance Computation*: After performing both the spelling and case ending modification for both the query word and searched words, Levenshtein distance LDM is computed for different modified pairs. Pairs containing the similar words result lower distance and vice versa. Since, this time edit distance is computed for the words being in modified state, words which are not exact match of the original query word may result zero edit distance (e.g. 'ঢাকা', 'ঢাকার', 'ঢাকাকে', 'ঢাকাতে', 'ঢাকায়' have zero LDM with 'ঢাকা'). As our goal is to sort the words in accordance with their phonetic and semantic similarity with the query word, the coarse edit distance computed earlier is also considered. This ensures exact matching words to have lower distance. For serving the purpose, a weighted sum of coarse distance CD and Levenshtein distance LDM for modified word is used. Thus, final distance FD after both phonetic and semantic consideration is-

$$F_D = C_D + k_w L_{DM} \quad (4)$$

Here, the term k_w is used to manipulate the result by giving higher or lower weight to any particular distance. The first term C_D was computed with no alteration in any of the words. So, giving lower value of k_w will result in a higher importance to the distance computed with no modification at all. On the other hand, higher value of k_w will cause the phonetic and semantic similarity have greater dominance on the result.

As we want to give both the distances to have almost equal priority, value of k_w is chosen on to do so. A closer observation on the equation used to compute C_D and that for computing the Levenshtein distance, reveals that, value of C_D is always higher than L_{DM} for the same type of mismatch. It was empirically found that, setting value of k_w to 4 gives us a satisfactory result.

Higher weight is given to L_{DM} due to the fact that, the evaluation procedure of coarse distance C_D is such that, it gives higher numerical value than that found from Levenshtein distance, computed over the same pair of words.

Finally, F_D will be considered as the distance between the query word and any word from the wordlist.

IV. OBSERVATION

For observing the performance of our work, we ran a search with 'মন্ত্রী' as the query word. Here, a deliberate spelling mistake was introduced ('মন্ত্রি' instead of 'মন্ত্রী' meaning minister) to examine the performance. A group of several words were taken to perform the search on. To assess the performance of our method, we compared the searched result found by our method, with that found by implement

tation of only Levenshtein distance without any modification made. In the search, threshold for our proposed method was used four times than that of used with pure Levenshtein distance. This is due to the computation procedure of distance in our method. In our method, small dissimilarity results a higher numerical value of distance.

In table IV, observation of order of searched words according to F_D reveals that, the word 'মন্ত্রী' (meaning 'minister') has the lowest distance and so first in the list. Both semantically and phonetically, 'মন্ত্রী' seems to be the best match here. Second and third on the list are 'মন্ত্রীর' (meaning 'minister's') and 'মন্ত্রীকে' (meaning 'to minister'). These two words are not phonetically well matched to the query word but from semantic standpoint it is obviously of greater interest than the other words in the database (মন্ত্র, মন্ত্রণ, তন্ত্রী, যন্ত্রী, তন্ত্রী, যন্ত্রী) which have no relevance with our word of interest here.

On the right side of the table, words are arranged in accordance with there Levenshtein distance with the query word. As stated earlier, this distance matching technique was not built to take the pros and cons of Bangla grammar and spelling, and thus unable to arrange the words as they are expected. A single substitution will convert 'মন্ত্রি' to মন্ত্রণ or মন্ত্রী, or a single deletion will convert it to মন্ত্র, resulting all these words to be treated with equal importance. Moreover, semantically similar words are send away later on the list due to their higher mismatch with the query word.

TABLE IV. PERFORMANCE COMPARISON

Query word মন্ত্রী			
Searched words sorted according to F_D	F_D	Searched words sorted according to only Levenshtein distance	L_{DM}
মন্ত্রী (minister)	2	মন্ত্রণ মন্ত্রী(minister) মন্ত্র	1
মন্ত্রীর (minister's)	3	মন্ত্রীর(minister's) মন্ত্রণা তন্ত্রী তন্ত্র যন্ত্র যন্ত্রী	2
মন্ত্রীকে (to minister)	4	মন্ত্রীকে (to minister) তান্ত্রিক যন্ত্রণা যান্ত্রিক	3
মন্ত্র	5	দন্ত দন্তী	4
মন্ত্রণ	6		
তন্ত্রী যন্ত্রী	8		
মন্ত্রণা তন্ত্র যন্ত্র	11		
তান্ত্রিক যান্ত্রিক	16		

To measure the running time of the process, a test was performed with different words as query words, and with different database files to search on. Articles from different online Bangla newspapers were used as database. For ease of understanding, a simple flow diagram of the string search process is shown in Figure 1. In the figure, between two consecutive stages, number of words meeting the criteria to go to the latter stage is shown. For example, on the first stage of filtering (Length Matching) N words are inserted. In the filtering process, large part of the words are discarded as they do not satisfy the criteria to pass through and only N_L words are passed to the next stage. Similarly, N_C words are allowed to pass through the second filter. All these words are then taken for modification and go through further processes. After finding the L_{DM} , words are sorted according their distances computed.

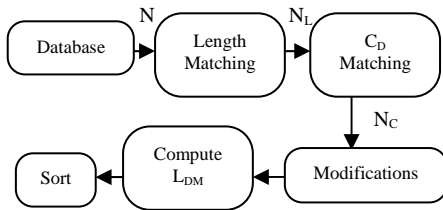


Figure 1. Simple flow diagram of the string search process

TABLE V. RUNNING TIME COMPARISON

Length of query word	N	N_L	N_C	Running time T_R (Second)	T_R (without fast filtering, only by using Levenshtein distance) (Second)
3	7476	2885	122	0.281	0.594
8	7476	6671	234	0.469	0.953
17	7476	1597	195	0.297	1.610
4	15947	9117	256	0.719	1.375
5	15947	12768	230	0.828	1.578
11	15947	9496	278	0.766	2.407
17	15947	3177	280	0.547	3.235

From Table V, it can be seen that, running time decreases when length of the query word is too high or too low. This is due to the fact that, most of the words from database cannot pass through the first stage, i.e. length matching (low value of N_L). It is found that, for all the cases, running time is always lower than finding the similarity solely by Levenshtein distance.

V. CONCLUSIONS

In this work we present a novel technique to implement approximate string search in Bangla. The spelling error modification on the first stage can be proved useful for dictionary search or spelling suggestion. Addition of clever rejection of inflections ensures words with related meanings to come up on the queue on web query or database search. The technique adopted in our work can be used for other

inflectional languages like Bangla provided that the inflection mechanism is well studied.

REFERENCES

- [1] A. Arasu, V. Ganti, and R. Kaushik, "Efficient Exact Set-Similarity Joins," in *VLDB*, 2006, pp. 918–929.
- [2] S. Chaudhuri, K. Ganjam, V. Ganti, and R. Motwani, "Robust and Efficient Fuzzy Match for Online Data Cleaning," in *SIGMOD*, 2003, pp. 313–324.
- [3] L. Gravano, P. G. Ipeirotis, H. V. Jagadish, N. Koudas, S. Muthukrishnan, and D. Srivastava, "Approximate string joins in a database (almost) for free," in *VLDB*, 2001, pp. 491–500.
- [4] E. Sutinen and J. Tarhio, "On Using q-Grams Locations in Approximate String Matching," in *ESA*, 1995, pp. 327–340.
- [5] R. Bayardo, Y. Ma, and R. Srikant, "Scaling up all-pairs similarity search," in *WWW Conference*, 2007.
- [6] E. Ukkonen, "Approximate String Matching with q-Grams and Maximal Matching," *Theor. Comput. Sci.*, vol. 1, pp. 191–211, 1992.
- [7] V. Levenshtein, "Binary Codes Capable of Correcting Spurious Insertions and Deletions of Ones," *Profl. Inf. Transmission*, vol. 1, p. 8–17, 1965.
- [8] S. Sarawagi and A. Kirpal, "Efficient set joins on similarity predicate," in *ACM SIGMOD*, 2004.
- [9] G.K. Saha, A. B. Saha and S. Debnath, "Computer Assisted Bangla Words POS Tagging," in Proc. International Symposium on Machine Translation NLP & TSS (iSTRANS-2004), 2004.
- [10] C. Barr, R. Jones and M. Regelson, "The Linguistic Structure of EnglishWeb-Search Queries," in Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, 2008, pp. 1021–1030.
- [11] The Unicode Consortium, The Unicode Standard, Version 4.0, Addison-Wesley, 2003. Also available online at <http://www.unicode.org/charts/PDF/U0980.pdf>.
- [12] A. K. Guha, *Notun Bangla Rachana*, 1st ed., Mullick Bros, Dhaka, Bangladesh, 1994.
- [13] E. Ukkonen, 'Approximate string-matching with n-grams and maximal matches', *Theoretical Computer Science*, 92, 191–211, (1992).
- [14] B. Ma, J. Tromp, and M. Li. PatternHunter: Faster and more sensitive homology search. *Bioinformatics*, 18:440–445, 2002.
- [15] K. M. Odell, and R. C. Russell, "Soundex phonetic comparison system" cf. U.S. Patents 1261167 (1918), 1435663 (1922).
- [16] T. N. Gadd, "PHONIX: The Algorithm", *Program*, 24(4), pp. 363–366, 1990.

Multicast Routing and Wavelength Assignment for Capacity Improvement in Wavelength Division Multiplexing Networks

N.Kaliammal

Professor, Department of ECE,
N.P.R college of Engineering and Technology,
Dindugul, Tamil nadu

Tel: +91 9965557267

G.Gurusamy

Prof/Dean/ EEE, FIE
Bannari amman Institute of Technology,
Sathyamangalam, Tamil nadu.

Tel: +91 9791301662

Abstract—In WDM network, the route decision and wavelength assignment of light-path connections are based mainly on the routing and wavelength assignment (RWA). The multicast routing and wavelength assignment (MC-RWA) problem is for maximizing the number of multicast groups admitted or for minimizing the call blocking probability. In this paper, The design of multicast routing and wavelength assignment technique for capacity improvement in wavelength division multiplexing (WDM) networks is proposed. In this technique, the incoming traffic is sent from the multicast source to a set of intermediate junction nodes and then, from the junction nodes to the final destinations. The traffic is distributed to the junction nodes in predetermined proportions that depend on the capacities of intermediate nodes. Then, paths from source node to each of the destination nodes and the potential paths are divided into fragments by the junction nodes and these junction nodes have the wavelength conversion capability. By using the concept of fragmentation and grouping, the proposed scheme can be generally applied for the wavelength assignment of multicast in WDM network. By simulation results, it is proved that the proposed technique achieves higher throughput and bandwidth utilization with reduced delay.

I. INTRODUCTION

A. Wavelength-Division-Multiplexing (WDM) Networks

The need for on-demand provisioning of wavelength routed channels with service differentiated offerings within the transport layer has become more essential due to the recent emergence of high bit rate IP network applications. Diverse optical transport network architectures have been proposed in order to achieve the above requirements. This approach is determined by the fundamental advances in the wavelength division multiplexing (WDM) technologies. Due to the availability of ultra long-reach transport and all-optical switching, the deployment of all-optical networks has been made possible [1].

The concurrent transmission of multiple streams of data with the assistance of special properties of fiber optics is called as wavelength division multiplexing (WDM). The WDM network provides the capability of transferring huge

amount of data at high speeds by the users over large distance [2].

For the future generation internet, WDM is considered as a backbone which is the most talented technology. The data is routed through optical channels called light paths in WDM all optical networks. The light path establishment requires same wavelength and it should be used along the entire route of the light path without wavelength conversion. This is commonly considered to the wavelength continuity constraint [3].

B. Multicasting in WDM Networks

A network technology which is used for the delivery of information to a group of destinations is called as multicast addressing. This simultaneously uses the most efficient strategy to deliver the message over each link of the network only once. Moreover, it creates the copies only when the links to the multiple destinations split [4].

In recent years, multicast communication is turning out to be vital due to its efficient resources usage and the increasing popularity of the point-to-multipoint multimedia applications. Usually, a source and a set of destinations are included in a multicast session. In conventional data networks, in order to allow a multicast session, a multicast tree which is rooted at the source is constructed with branches spanning all the destinations [5].

Recently, multicast routing in optical networks has been researched which is related to the design of multicast-capable optical switches. For multicast in WDM networks, the concept of light-trees was introduced. Reducing the distance of network-wide hop and the total number of transceivers used in the network are the objective of setting up the light trees. Nowadays, there are several network applications which require the support of QoS multicast such as multimedia conferencing systems, video on demand systems, real-time control systems, etc. [6].

C. Routing and Wavelength in WDM

In WDM network, the route decision and wavelength assignment of light-path connections are based mainly on the routing and wavelength assignment (RWA). This is the most

important and basic issue in resource management. For maximizing the number of multicast groups admitted or for minimizing the call blocking probability with certain number of wavelengths, the multicast routing and wavelength assignment (MC-RWA) problem is studied. [7].

The problem of finding a multicast tree and allocating available wavelength for each link of the tree is known as the Multicast Routing and Wavelength Assignment (MC-RWA) problem, which plays a key role in supporting multicasting over WDM networks [8]. The problems involving in the routing and wavelength assignment in WDM are as follows:

- Improper wavelength assignment, especially for the multicast connection, will cause wavelength blocking, whereas the network resources may be still underutilized.
- The wavelength continuity constraint, *i.e.*, that links from source to destination shall use the same wavelength to convey data in the same lightpath, always makes the wavelength assignment inflexible and causes wavelength blocking.
- The available wavelength can be maximized by the wavelength converter but this type of device is much intricate and cost is also high when compared with the type of device which cannot perform the conversion.
- The signal may also decay during the conversion. Therefore, it is not possible to have all network nodes be equipped with wavelength conversion capability.
- The problem of the node architecture is that they were designed without having into account power efficiency, neither complexity of fabrication [9].
- The two sub-problems of the routing and wavelength assignment are the routing problem and the wavelength assignment problem, which can be either coupled or uncoupled. In the case of uncoupled situation, initially a route or a tree is obtained which is then followed by the wavelength assignment where the trees must be kept unchanged and is called as the static RWA. In the coupled case, based on the state of the wavelength assignment, the routes are decided which is usually called as dynamic or adaptive RWA [7].

In previous paper, a resource efficient multicast routing protocol is developed. In this protocol, the incoming traffic is sent from the multicast source to a set of intermediate junction nodes and then, from the junction nodes to the final destinations. The traffic is distributed to the junction nodes in predetermined proportions that depend on the capacities of intermediate nodes. Bandwidth required for these paths depends on the ingress–egress capacities, and the traffic split ratios. The traffic split ratio is determined by the arrival rate of ingress traffic and the capacity of intermediate junction nodes [13].

In this paper, a multicast routing and wavelength assignment technique in wavelength division multiplexing networks is designed. In this technique, paths from source node to each of the destination nodes and the potential paths

are divided into fragments by the junction nodes and these junction nodes have the wavelength conversion capability. By using the concept of fragmentation and grouping, the proposed scheme can be generally applied for the wavelength assignment of multicast in WDM network. The Least Influence Group (LIG) approach is used to provide the wavelength selection.

II. RELATED WORK

Jingyi He et al [7] have proposed for the first time a formulation of the MC-RWA problem with the objective to maximize the number of multicast groups admitted, or equivalently, to minimize the call (or session) blocking probability given a certain number of wavelengths. The formulation is a nonlinear integer program, which in general is complex to solve so a near-optimal solution of the problem is proposed using a two-step approach based on linear programming. The drawback in this work is that the focus is on minimizing the user blocking probability instead of the session blocking probability for single-source applications.

Anping Wang et al [8] have proposed a new multicast wavelength assignment algorithm called NGWA with complexity of $O(N)$, where N is the number of nodes on a multicast tree. The whole procedure of NGWA algorithm is separated into two phases: the partial wavelength assignment phase and the complete wavelength assignment phase. The drawback of this work is that this method achieves only satisfactory performance in terms of the total number of wavelength conversions and the average blocking probability

Nina Skorin-Kapov [10] has addressed the problem of multicast routing and wavelength assignment (MC RWA) in wavelength routed WDM optical networks. Multicast requests are facilitated in WDM networks by setting up so-called light-trees and assigning wavelengths to them. She has proposed a heuristic algorithm based on bin packing methods for the general MC RWA problem, which is NP-complete. These algorithms can consider unicast, multicast and broadcast requests with or without QoS demands. Computational tests indicate that these algorithms are very efficient, particularly for dense networks.

Fen Zhou et al [11] have proposed a routing and wavelength assignment for supporting multicast traffic is investigated in WDM mesh networks under sparse splitting constrain. This problem is generally solved in two phases respectively with the purpose of minimizing the number of wavelengths required. Alternative routing is first proposed to route each session by pre-computing a set of candidate light-forests. Then wavelength assignment is formulated as coloring problems by constructing a conflict graph. Potential heuristic algorithms are proposed. The drawback of this work is that simulation should be done to assess the verification of the proposed methods.

Yuan Cao et al [12] have proposed an efficient QoS-guaranteed Group Multicast RWA solutions, where the transmission delay from any source to any destination within a multicast group is within a given bound. They have formulated the QoS-guaranteed GMC-RWA problem as an in-group traffic grooming and multicasting problem, where traffic

streams from members of the same group are groomed in an effective way before being delivered to their common destinations, subject to the following optical layer constraints.

III. MULTICAST TREE FORMATION

A. Basic Definitions

The node which cannot split the incoming message to the outgoing ports is called as Multicast Incapable (MI) nodes. But it can utilize a small amount of optical power from the wavelength channel while forwarding it to only one output link.

The nodes which are capable of splitting the incoming message to all the outgoing ports are called as Multicast Capable (MC) nodes.

The set which includes the multicast capable nodes (MC node) and the leaf multicast incapable nodes (leaf MI nodes) is called as MC_SET.

The set which includes only the non-leaf multicast incapable nodes, which are not able to connect a new destination to the multicast tree, is called as MI_SET.

The set D includes the unvisited multicast destinations which are not yet joined to the multicast tree.

A constraint path between a node u and a tree T is a shortest path from node u to a node v in the MC_SET for T , and this shortest path should not traverse any node in MI_SET for T . And the constraint path with the minimum length is called the Shortest Constraint Path (SCP).

For one nearest destination d , MC_SET may have different SCPs to the sub-tree. Let X and Y are the nodes for the sub-tree in MC_SET. Without involving any node in MI_SET for the sub-tree, both the shortest paths from X and Y to the nearest destination d have the shortest length among all the nodes in MC_SET. Here, the nodes like X and Y are named as junction nodes in the sub-tree.

Member only Algorithm

```

T = {s}
MI_SET = Null
MC_SET = {s}
D = {D1, D2, ..., Dn}
1. For each Di, where i = 1, 2, ..., n
    1.1 If dist (Di, N) = min, where N ∈ MC_SET, then
        1.1.1 Add Di to T
        1.1.2 Find SCP (Di, T) ∉ M, where M ∈ MI_SET
        1.1.3 Add SCP (Di, T) to T
        1.1.4 Add all the MC nodes to MC_SET
        1.1.5 Add all the leaf MI nodes to MC_SET
        1.1.6 Add all the non-leaf MI nodes to MI_SET
        1.1.7 Delete the non - leaf MI node from MC_SET
        1.1.8 Delete the destination di from

```

D.

1.2 End if

2. End for

3. If {D} ≠ Null, Then

3.1 Repeat from 1.

4. End if

B. Multicast Routing

A collection of point to multiple point paths from the source node to each destination is considered as a multicast tree. Choosing a suitable wavelength for its downlink is flexible for a path in the WDM network which has sparse junction nodes. The main objective is to reduce the affected capacity. This can be done by selecting a suitable wavelength for the downlink of the junction nodes which reduces the influence on the potential request paths across it. The junction node is considered as an end point of a wavelength within a fragment. According to the position of converters within the path, the path can be divided into uni-wavelength fragments. As a result, paths from source node to each of the destination nodes and the potential paths are divided into fragments by the junction nodes and these junction nodes have the wavelength conversion capability.

A network $G = (N, E)$ with node set N and (directed) edge E set is taken, where each node in the network can be a source or destination of traffic. The nodes in N are $\{N1, N2, \dots, Nn\}$.

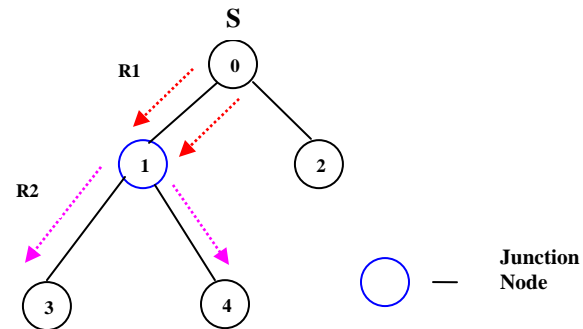


Figure 1. Multicast Routing Process

The above diagram (Fig. 1) shows the routing process. A predetermined fraction of the traffic entering the network at any node is distributed to every junction node. The corresponding route from the source to the junction node can be denoted as $R1$. Then each junction node receives the traffic to be transmitted for different destinations and it routes to their respective destinations. The corresponding route from the junction node to the destination can be denoted as $R2$.

Let I_i and E_i be the constraints on the total amount of traffic at ingress and egress nodes of the network, respectively.

The traffic along $R1$ and $R2$ must be routed along bandwidth-guaranteed paths. Bandwidth required for these paths depends on the ingress-egress capacities, and the traffic split ratios. The traffic split ratio (δ) is determined by the arrival rate of ingress traffic and the capacity of intermediate junction nodes.

The bandwidth requirement for the routing paths R1 and R2 is derived. Consider a node i with maximum incoming traffic I_i . Node i sends $\delta_i I_i$ amount of this traffic to node j during R1 routing for each $j \in N$ and thus the traffic demand is $\delta_i I_i$. Now, node i has received $\delta_i I_k$ traffic from any other node k . Out of this, the traffic destined for node j is $\delta_i I_{kj}$ since all traffic is initially split without regard to the final destination. The traffic that needs to be routed from node i to node j at R2 routing is given below:

$$\sum_{k \in N} \delta_i r_{kj} \leq \delta_i E_j.$$

Thus, the traffic demands from node i to node j at the end of R2 routing is $\lambda_i E_j$.

Hence, the maximum demand from node i to node j as a result of R1 and R2 routing is $\delta_i I_i + \delta_i E_j$.

Let $M = [m_{ij}] = [\delta_i I_i + \delta_i E_j]$ be the fixed matrix which can handle the traffic variation. It depends only on aggregate ingress-egress capacities and the traffic split ratios $\delta_1, \delta_2, \dots, \delta_n$. Thus the routing scheme is unaware to the changes in traffic distribution.

IV. MULTICAST WAVELENGTH ASSIGNMENT

A. Grouping the Paths

Assume the set $R_i = \{R_i^1, R_i^2, \dots, R_i^j, \dots\}$ to represent all fragments of the path from source to the i^{th} destination in the multicast tree. R_i^j is the j^{th} fragment of the path i . If AWR_i^j is the set of available wavelengths of the j^{th} fragment of path i , then the number of wavelengths in AWR_i^j is regarded as the capacity of this fragment. The capacity of the j^{th} fragment of the path i , $SCPR_i^j$ is obtained as

$$SCPR_i^j = \begin{cases} OL(S, J_i^k) = |AWR_i^j|, & k=1, j=1 \\ OL(J_i^k, J_i^{k-1}) = |AWR_i^j|, & 1 < k \leq M_i, j=k \\ OL(J_i^k, D) = |AWR_i^j|, & k=M_i, j=M_i+1 \end{cases} \quad (1)$$

where s is the source node of the multicast tree, D_i is the i^{th} destination of the multicast tree, J_i^k is the k^{th} wavelength converter in the path i , and $M_i + 1$ is the number of fragments of path i if there are M_i junction nodes being traveled by the path. The Overlap function $OL(n1, n2)$ represents the size of the intersection set of all available wavelengths for all links from node $n1$ to $n2$.

For the potential request paths, the set $P_i = \{p_i^1, p_i^2, \dots\}$ is defined to indicate all fragments of the i^{th} potential request path and the capacity of the j^{th} fragment of the potential path i , $SCPP_i^j$, can be stated as following

$$SCPP_i^j = \begin{cases} OL(S, J_i^k), & k=1, j=1 \\ OL(J_i^k, J_i^{k-1}), & 1 < k \leq M_i, j=k \\ OL(J_i^k, D), & k=M_i, j=M_i+1 \end{cases} \quad (2)$$

where P_i^j is the j^{th} fragment of the potential path i , S_i is the source node of the potential path i , and D_i is the destination of the potential path i . Basically, each fragment can be treated as a reassignment domain of wavelength.

Fragments of a path are mutually independent from the wavelength assignment point of view and may be with different fragment capacities. The actual capacity of a path is basically determined by its fragment(s) with the least capacity. The fragment(s) with the least capacity of a path is named the critical fragment of that path. Let CP_i and CPP_i be the path capacity (the least fragment capacity) of the path i of the multicast tree, and the path capacity of the potential path i , respectively, then

$$CP_i = \min_{1 \leq j \leq r_i+1} SCPR_i^j \quad (3)$$

and

$$CP_{pi} = \min_{1 \leq j \leq r_i+1} SCPP_i^j \quad (4)$$

Capacity of the path cannot be decreased by decreasing the capacity in a fragment whose capacity is larger than the critical fragment of that path. A path may have more than one critical fragment. Let $F_i = \{f_i^1, f_i^2, \dots\}$ be the set of the critical fragments in the potential path i . Then F_i can be used to indicate whether the potential path is affected or not during the wavelength assignment of the multicast tree. So, the critical fragment of a potential path is the fragment traveled by the multicast tree. The impact on the potential path can be reduced by considering the wavelength assignment of that fragment carefully. Fragments which come from multicast tree with common links into groups are coupled using the concept of grouping. Within a group, all fragments have common wavelengths. As a result a group is composed of fragments whose links are overlapped.

$$G = \{G1, G2, \dots, G_m, \dots, GY\} \quad (5)$$

where G is the set of all groups in a multicast tree, G_m is the set of all fragments in the m^{th} group.

The multicast tree with n destinations is treated as n unicast paths from source to each destination. Paths are fragmented with respect to junction nodes. Same group fragments have more than one available wavelength in common. Let AWG_m be the connection set of all fragments existing wavelengths in the m^{th} group. The group capacity, CG_m , is defined as the number of wavelengths in AWG_m . If links of a fragment and the links in the m^{th} group are overlapped and no common available wavelength between them, this fragment will be considered as a new group.

B. Total Network Capacity Estimation

The influence of network capacity is examined by checking whether the links of potential paths overlap with those of the multicast groups. If the overlap occurs at the critical fragments of the potential path and the assigned wavelength is the one of the available wavelengths in that critical fragment, the path capacity of the potential path will be affected. Let $C_m(p_i, \lambda)$ be the capacity of p_i being influenced when the wavelength λ is assigned in the m^{th} group, and x be a

common link of the m^{th} group and the critical fragment of the potential path i . Then

$$C_m(p_i, \lambda) = \begin{cases} 1 & \text{if } (\lambda \in x_w) \wedge x \in LS_{m, Fi} \\ 0 & \text{Otherwise} \end{cases} \quad (6)$$

where $LS_{m, Fi} = L_{Gm} \cap L_{Fi}$, L_{Gm} is the set of all links in the m^{th} group, L_{Fi} is the set of all links in the critical fragments of the potential path i . and x_w is the set of all available wavelengths on link x .

The network capacity affected when λ is assigned for the m^{th} group, $TNC_{m, \lambda}$, can be obtained by the summation of the influence of all potential paths as

$$TNC_{m, \lambda} = \sum_{p_i \in P} C_m(p_i, \lambda) \quad (7)$$

The total network capacity (TNC) gets affected since each group should assign one wavelength, and it can be obtained by the summation as

$$TNC = \sum_{All\ m} \sum_{p_i \in P} C_m(p_i, \lambda_m) - q \quad (8)$$

In the m^{th} group, λ_m is the wavelength assigned and q is the affected capacity that is counted repeatedly. This can be regained in the first term of (8). When the same wavelength is assigned to the groups it leads to repeated counts and also the critical fragments of the path travels through the group. For example, the available wavelengths of the critical fragment of potential path p_1 are (λ_1, λ_2) . G_1 and G_2 are the groups of the multicast tree. If λ_1 is assigned to G_1 and G_2 and if the critical fragment of potential path p_1 travels through G_1 and G_2 , then, according to the first term of (8), the affected capacity of p_1 is calculated twice. In fact, the decreased capacity is only one. The other repeated count happens when the same or a different wavelength is assigned to the groups and more than one critical fragment of an individual path goes through these groups.

C. Wavelength Assignment

By using junction nodes the multicast tree is separated into groups, so the wavelength assignments for groups are independent of each other. The wavelength assigned in the previous group has no effect on the wavelength assigned in the current group. The wavelength assigned for each group can be easily selected since all of the available wavelengths for a group have been collected in AWG_m .

The Least Influence Group (LIG) algorithm selects the wavelengths for groups to maximize the network capacity. The idea behind LIG algorithm is that the wavelength having the least effect on the potential paths is chosen for that group. The affected network capacity in (7) examines the influence of each wavelength assignment. The LIG algorithm is illustrated below:

$$AWG_m = \{\lambda_1, \lambda_2, \lambda_3, \dots\}$$

1. Find all p_b whose links overlap in the links of group m
2. For each $\lambda \in AWG_m$

$$TNC_{m, \lambda} = \sum_{p_b \in P'} C_m(p_b, \lambda)$$

3. Assign λ which $TNC_{m, \lambda}$ is minimum in group m

V. SIMULATION RESULTS

A. Simulation Model and Parameters

In this section, the performance of multicast routing and wavelength assignment technique is simulated with an extensive simulation study based upon the ns-2 network simulator [14]. The Optical WDM network simulator (OWNs) patch in ns2 is used to simulate a NSF network (Fig.1) of 14 nodes. Various simulation parameters are given in table I.

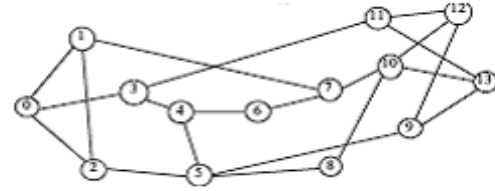


Figure 1. NSF network of 14 nodes

TABLE I: SIMULATION PARAMETERS

Topology	Mesh
Total no. of nodes	14
Link Wavelength Number	8
Link Delay	10ms
Wavelength Conversion Factor	1
Wavelength Conversion Distance	8
Wavelength Conversion Time	0.024
Link Utilization sample Interval	0.5
Traffic Arrival Rate	0.5
Traffic Holding Time	0.2
Packet Size	200
No. of Receivers	4
Max Requests Number	50
Rate	2,4,6, 8 and 10 Mb
Number of Traffic Sources	1,2,3,4 and 5

In this simulation, a dynamic traffic model is used, in which connection requests arrive at the network according to an exponential process with an arrival rate r (call/seconds). The session holding time is exponentially distributed with mean holding time s (seconds).

The connection requests are distributed randomly on all the network nodes. In all the simulation, the results of MRWA with the previous paper "resource efficient multicast routing (REMR) protocol [13]." Is compared.

B. Performance Metrics

In this simulation the blocking probability, end-to-end delay and throughput is measured.

- **Bandwidth Utilization:** It is the ratio of bandwidth received into total available bandwidth for a traffic flow.
- **Average end-to-end delay:** The end-to-end-delay is averaged over all surviving data packets from the sources to the destinations.
- **Throughput:** It is the number of packets received successfully.

C. Results

A. Effects of Varying Traffic

In the initial simulation, the traffic rate is varied as 2Mb, 4Mb, 6Mb, 8Mb and 10Mb and measure the throughput, end-to-end delay and bandwidth utilization.

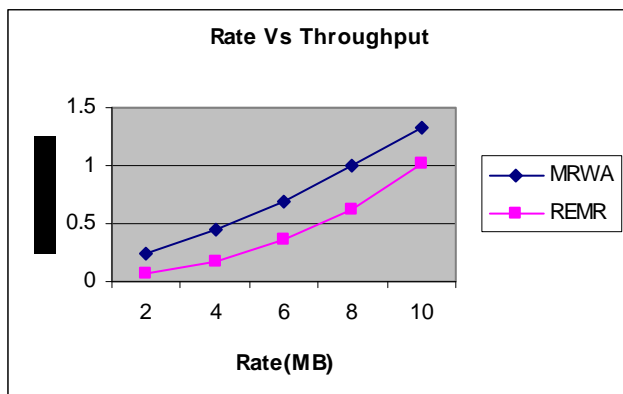


Figure 2. Rate Vs Throughput

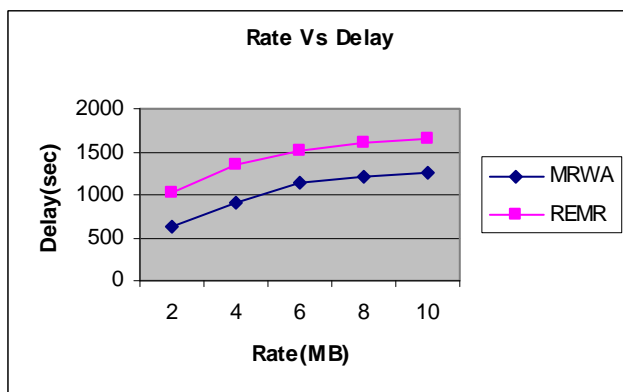


Figure. Rate Vs Delay

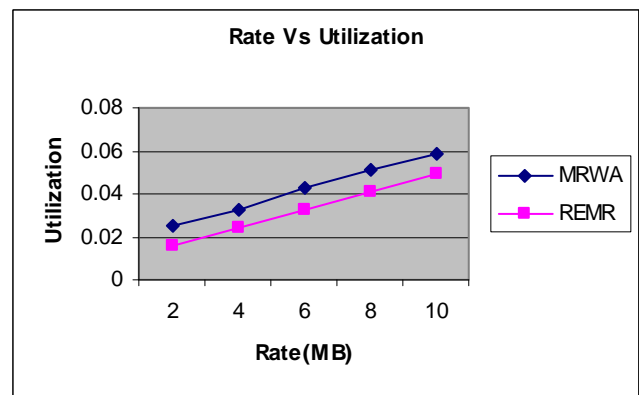


Figure 4. Rate Vs Utilization

Figure. 2 shows the throughput occurred when the rate is increased. From the figure, it is proved that the throughput is more in the case of MRWA when compared to REMR.

Figure.3 shows the end-to-end delay occurred when the rate is increased. It shows that the delay of MRWA is significantly less than the REMR.

Figure.4 shows the bandwidth utilization obtained when the rate is increased. MRWA shows better utilization than the REMR scheme.

B. Effect of Varying Traffic

In this simulation, the number of traffic sources is varied as 1, 2, 3, 4 and 5 and measure the throughput, end-to-end delay and bandwidth utilization.

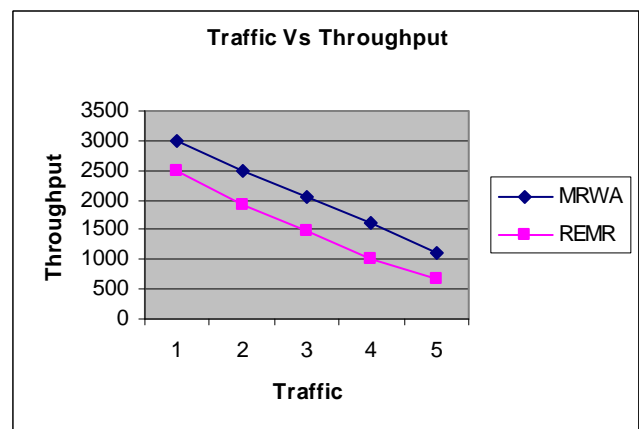


Figure 5. Traffic Vs Throughput

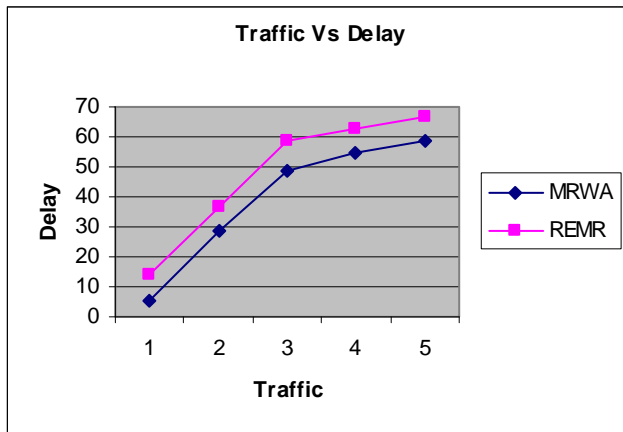


Figure 6. Traffic Vs Delay

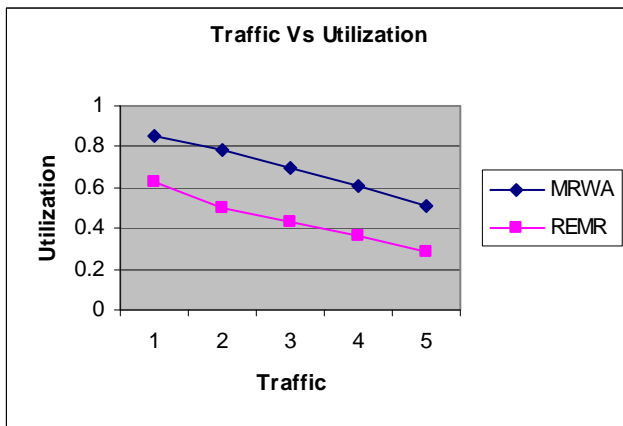


Figure 7. Traffic Vs Utilization

Figure 5 shows the throughput occurred when varying the number of traffic sources. From the figure it is proved that, the throughput is more in the case of MRWA when compared to REMR.

Figure.6 shows the end-to-end delay occurred when varying the number of traffic sources. It shows that the delay of MRWA is significantly less than the REMR.

Figure 7 shows the bandwidth utilization obtained when varying the number of traffic sources. MRWA shows better utilization than the REMR scheme.

Name of performance metrices	Effects on varying rate		Effects on varying Traffic	
	REMR	MRWA	REMR	MRWA
THROUGHPUT	0.49	0.71	1520	2000
DELAY	1440sec	1020sec	48.6sec	39.2sec
UTILISATION	0.0325	0.041	0.456	0.7

VI. CONCLUSION

In this paper, a multicast routing and wavelength assignment technique in WDM networks is developed. In this technique, the incoming traffic is sent from the multicast source to a set of intermediate junction nodes and then, from the junction nodes to the final destinations. The traffic is

distributed to the junction nodes in predetermined proportions that depend on the capacities of intermediate nodes. Then, paths from source node to each of the destination nodes and the potential paths are divided into fragments by the junction nodes and these junction nodes have the wavelength conversion capability. In order to select the wavelengths for groups to maximize the network capacity, the Least Influence Group (LIG) algorithm is used, i.e. the wavelength having the least effect on the potential paths is chosen for that group. So the affected network capacity influences the wavelength assignment. By simulation results, it is proved that the proposed technique achieves higher throughput(22% increase) and bandwidth utilization (1% increase)with reduced delay(420sec decrease)for varying rate & 9.4 sec decrease in delay ,0.3% increase in utilization and 480times increase in throughput for varying traffic.

REFERENCE

- [1] A. Rajkumar and N.S.Murthy Sharma, "A Distributed Priority Based Routing Algorithm for Dynamic Traffic in Survivable WDM Networks", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.11, November 2008.
- [2] Canhui (Sam) Ou Hui Zang, Narendra K. Singhal, Keyao Zhu, Laxman H. Sahasrabudhe, Robert A. Macdonald and Biswanath Mukherjee, "Sub path Protection For Scalability And Fast Recovery In Optical WDM Mesh Networks", IEEE Journal On Selected Areas In Communications, Vol. 22, No. 9, November 2004.
- [3] Vinh Trong Le, Son Hong Ngo, Xiao Hong Jiang, Susumu Horiguchi and Yasushi Inoguchi, "A Hybrid Algorithm for Dynamic Lightpath Protection in Survivable WDM Optical Networks", IEEE, 2005.
- [4] Multicasting: <http://en.wikipedia.org/wiki/Multicasting>
- [5] Fen Zhou, Miklos Molnar and Bernard Cousin, "Distance Priority Based Multicast Routing in WDM Networks Considering Sparse Light Splitting", IEEE 11th Singapore International Conference on Communication Systems – 2008
- [6] Xiao-Hua Jia, Ding-Zhu Du, Xiao-Dong Hu, Man-Kei Lee, and Jun Gu, "Optimization of Wavelength Assignment for QoS Multicast in WDM Networks", IEEE Transactions on Communications, Vol. 49, No. 2, February 2001.
- [7] Jingyi He, S.H. Gary Chan and Danny H.K. Tsang, "Routing and Wavelength Assignment for WDM Multicast Networks", In the proceedings of the IEEE GLOBECOM 2001.
- [8] Anping Wang, Qiwu Wu, Xianwei Zhou and Jianping Wang, "A New Multicast Wavelength Assignment Algorithm in Wavelength-Converted Optical Networks", International Journal of Communications, Network and System Sciences, 2009
- [9] G. M. Fernandez and D. Larrabeiti, "Contributions for All-Optical Multicast Routing in WDM Networks", 16th International Congress of Electrical Engineering, Electronics and Systems, IEEE INTERCON, 2009
- [10] Nina Skorin-Kapov, "Multicast Routing and Wavelength Assignment in WDM networks: A Bin Packing Approach", In the proceedings of optics Infobase in the optical networks, 2006.
- [11] Fen Zhou, Miklos Molnar and Bernard Cousin, "Multicast Routing and Wavelength Assignment in WDM Mesh Networks with Sparse Splitting", The 5th International Workshop on Traffic Management and Traffic Engineering for the Future Internet, Dec- 2009.
- [12] Yuan Cao and Oliver Yu, "QoS-Guaranteed Routing and Wavelength Assignment for Group Multicast in Optical WDM Networks", Conference on Optical Network Design and Modeling, 2005
- [13] N.Kaliyammal and G. Gurusamy, "Resource Efficient Multicast Routing Protocol for Dynamic Traffic in Optical WDM Networks", European Journal of Scientific Research, 2010
- [14] Network Simulator: www.isi.edu/nsnam/ns



N. Kaliammal received the B.E (ECE), M.E(Applied electronics), degrees from the Department of Electronics and Communication Engineering ,from Madurai Kamaraj University , Bharathiar University, Tamilnadu , in 1989, 1998, respectively. From 1990 to 1999, she served in the PSNA College of Engineering & Tech, Dindigul, Tamilnadu, as Lecturer. From 1999 to 2009, she was in RVS College of Engineering & Tech, Dindigul, Tamil Nadu, as assistant professor and Associate Professor. Currently she is working as Professor in NPR College of Engineering &Technology She is currently

pursuing the Ph.D. degree in Optical Networking, under the guidance of Dr.G.Gurusamy, Dean and Head/ EEE Department, Bannariamman Institute of Technology, Sathyamangalam, Tamil Nadu.

Dr.G.Gurusamy, received his BE, ME and PhD degree from PGS college of technology-Coimbatore. He has 35 years of teaching experience in PSG College of technology-Coimbatore. He is currently working as a Prof & Dean/in EEE Department of Bannariamman Institute of Technology-Sathyamangalam.

Blind Robust Transparent DCT-Based Digital Image Watermarking for Copyright Protection

Hanan Elazhary and Sawsan Morkos

Computers and Systems Department
Electronics Research Institute
Cairo, Egypt

Abstract—This paper proposes a novel technique for grayscale digital image watermarking using monochrome watermarks in the DCT domain. The proposed technique combines the contradictory goals of robustness against intentional and unintentional attacks with transparency in addition to blindness. We propose the combination of three techniques: Torus Automorphism (TA) permutation, a pseudo-random number generator, and visual cryptography to achieve high degree of robustness against intentional attacks. On the other hand, robustness against unintentional attacks is achieved not only by watermarking the low frequency DCT components, but also by intelligently utilizing the more robust relative values of these low frequencies instead of their absolute values. Experimental results using two test images prove the robustness of the proposed technique against several types of unintentional attacks: median filtering, blurring, sharpening, Gaussian noise addition, salt and pepper noise addition, and JPEG compression. Blindness is achieved using only few keys and two transparencies or shares for watermark extraction. In the literature, the size of the shares is usually double the size of the host image and consequently the process of generating the shares is usually time-consuming. In the proposed technique, generating the shares is straightforward and fast. Also, their size is exactly equal to the watermark size. To achieve 100% transparency, the proposed technique does not change any pixel values of the host image.

Keywords- watermarking; image processing; security

I. INTRODUCTION

The Internet has become the most popular channel for transmitting various forms of multimedia digital data. Multimedia data in digital format can be modified and illegally used with ease. Thus, the copyright protection of digital images transmitted over the Internet has become an important research topic in recent years. In the past few years, two research topics that are related to but different from the above topic have been proposed in the literature. One is image authentication and the other is image data hiding. The goal of image authentication is to verify the originality of an image by detecting malicious manipulations. Most of the earlier techniques in the literature treated some practical manipulations such as image compression and image enhancement as attacks. Lin and Chang proposed an image authentication technique which can detect malicious manipulations but allow JPEG lossy compression [1]. Hiding data in images involves embedding a large amount of secret data into the host image with minimal perceptible

degradation of host image quality. Unfortunately, the hiding capacity of secret data and the distortion of the host image are a tradeoff since more hidden data always result in more distortion of the host image [2].

Digital image watermarking is a method for ownership assertion and copyright protection. In such a technique, the watermark is embedded into the host image such that the embedded watermark can be later extracted to make an assertion about the host image ownership. There are two essential requirements for this purpose. The first requirement is the invisibility or transparency of the embedded watermark. In other words, the embedded watermark should not be perceived by human eyes and should not degrade the quality of the watermarked host image. The other requirement is the robustness of the embedded watermark. In other words, the embedded watermark should be able to resist both intentional and unintentional attacks. In intentional attacks, the attackers try to extract the embedded watermark for subsequent destruction. In other words, they alter some parameters of the watermarked host image to prevent legal owners from extracting the watermark for ownership assertion. In unintentional attacks, on the other hand, the watermarked host image is treated using image processing techniques including compression and filtering. Digital image watermarking could be done in the spatial domain [3-5] or in a transform domain including Discrete Cosine Transform (DCT) [6-8], Discrete Wavelet Transform (DWT) [9, 10] or a combination of such domains [11, 12]. Embedding a watermark into a transform domain rather than the spatial domain is more robust, and has higher resistance to various attacks [13]. Digital image watermarking techniques can be classified according to the watermark extraction process as follows [4]:

- Non-blind techniques that require both the secret key(s) for watermark embedding and the original host image.
- Semi-blind techniques that require both the secret key(s) for watermark embedding and the embedded watermark bit sequence.
- Blind techniques that require only the secret key(s) for watermark embedding. Neither the original host image nor the embedded watermark bit sequence is needed.

In this paper, we propose a novel digital-image watermarking technique that incorporates the important characteristics discussed above (robustness, transparency, and

blindness) for embedding a monochrome watermark into a grayscale host image. Since, as mentioned in Section I, it has been shown that embedding a watermark into a transform domain rather than the spatial domain is more robust, and has higher resistance to various attacks [13], the proposed technique embeds the watermark in the DCT domain. This technique is discussed in the rest of the paper.

The paper is organized as follows: Section II discusses related research in the literature and discusses the shortcomings that call for developing the technique proposed in the paper. Section III describes the proposed technique in terms of the embedding and the extraction processes and emphasizes its advantages. Section IV provides experimental results that demonstrate the robustness of the proposed technique against unintentional attacks. Finally, Section V provides the discussion and conclusions of the paper.

II. RELATED WORK

Many techniques in the literature have been proposed for digital image watermarking for copyright protection. It has been shown that embedding a watermark into a transform domain rather than the spatial domain is more robust and has higher resistance to attacks [13]. Thus, most of these techniques embedded the watermark in the transform domain(s) to achieve robustness against unintentional attacks. The DCT domain is one of these transform domains that has been frequently used in the literature. In the DCT domain, embedding the watermark in the insignificant components (high frequencies) makes hiding the watermark easier, but the embedded watermark is less resilient to attacks. On the other hand, embedding the watermark in the most significant components (DC or low frequencies) makes the embedded watermark robust against unintentional attacks. But, this makes hiding the watermark harder, which may degrade the host image quality. This is because low frequency components contain more energy than high frequency components and the human perception system is more sensitive to low frequency components. Saryazdi et al. [6] embedded a given watermark bit in the five first low AC coefficients. Similarly, Lin et al [7] embedded two watermark bits in the seven low frequency locations. Patra et al. [8] randomly selected a low frequency location for embedding a given watermark bit. However, the disadvantage of these techniques is that they physically embed the watermark bits which result in changing some pixels of the host image and degrading its quality. Also, another disadvantage of the above techniques is that they utilize the absolute values of the DCT coefficients. Unfortunately, these values can easily change under unintentional attacks. Thus, some techniques in the literature called for 100% transparent watermarking without physical embedding of the watermark [3, 4]. The result is a watermarked image with no distortion. The proposed technique does not change any pixel value of the DCT-transformed host image and also attempts to utilize the more robust relative values of the DCT coefficients instead of the absolute values.

Several research studies in the literature aimed at developing watermarking techniques that are robust against intentional attacks. This has been achieved by the use of Torus

Automorphism (TA) permutation for scrambling the watermark [7, 14-16] before embedding into the host image. Using a pseudo-random number generator to select arbitrary host image blocks and/or watermark bits for embedding has also been proposed in the literature [8]. The proposed technique combines Torus Automorphism (TA) permutation and a pseudo-random number generator to increase robustness against intentional attacks.

To increase this robustness even further, digital image watermarking based on visual cryptography [13], [17] has been proposed. Visual cryptography uses a codebook to divide a watermarked host image into several different transparencies, shares, or sharing images. When the copyright problem needs to be resolved, the ownership verification information can be obtained by stacking several sharing images. One or more of these shares could be registered to the Certified Authority (CA) for additional security and protection against intentional attacks. Because visual cryptography has good performance in protecting copyright of digital media [18], it has gained much attention of researchers in recent years. The problem with these techniques is that generating the shares is always time consuming. Besides, the size of each share is usually too large. For example, according to the Naor and Shamir scheme [17], each pixel of a given host image is replaced by 2×2 pixels. Hence, a host image with M by N pixels can be divided into two sharing images with $2M$ by $2N$ pixels. This scheme is the predominant technique used in the literature for this purpose [13]. This calls for developing a technique that can quickly generate shares of much smaller sizes. The idea of our proposed technique is to quickly generate a different kind of shares with much smaller size as long as they have the same advantage: helping in obtaining ownership verification information. The proposed technique is described in the following section.

0	1		
2	3		

Figure 1. Possible embedding locations in a 4×4 DCT-transformed block.

III. THE PROPOSED TECHNIQUE

The proposed technique can be described in terms of the watermark embedding process and the watermark extraction process.

A. The Watermark Embedding Process

The steps of the proposed watermark embedding process can be described as follows:

- **Step 1:** Divide the host image into non-overlapping 4×4 blocks.
- **Step 2:** Use Torus Automorphism (TA) permutation [14] to disarrange the watermark bits using the following equation:

$$\begin{pmatrix} i^* \\ j^* \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ k & k+1 \end{pmatrix} * \begin{pmatrix} i \\ j \end{pmatrix} \bmod m \quad (1)$$

Equation (1) indicates that each bit of the watermark at location (i, j) will be moved to a new location (i*, j*). Parameters k and m are secret keys needed for disarranging and rearranging the watermark bits. These permuted watermark bits will be embedded bit by bit and line by line into the host image.

- **Step 3:** Arrange the host image blocks sequentially one by one and row by row. Use a pseudo-random number generator to determine the sequence of host image blocks used for embedding the permuted watermark bits. A pseudo-random number generator with a given seed always generates the same sequence of random numbers. Thus, the seed of the pseudo-random number generator is a secret key needed to embed and extract the permuted watermark bits. Steps 2 and 3 offer cryptographic protection against intentional attacks because the secret keys of the pseudo-random number generator and TA permutation are necessary for extracting and rearranging the embedded watermark for subsequent destruction.
- **Step 4:** DCT-transform the next host image block from the spatial domain to the frequency domain. Each block is DCT-transformed using the following equation [6]:

$$DCT(i, j) = C(i) * C(j) * \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} pixel(x, y) * \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right] \quad (2)$$

where $C(i), C(j) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } i, j=0 \\ \sqrt{\frac{2}{N}} & \text{otherwise} \end{cases}$

In equation (2), DCT (i, j) represents the value at location (i, j) in the DCT-transformed block, while pixel (x, y) represents the value at location (x, y) in the original block. N is the number of locations (pixels) in each block.

- **Step 5:** Embed the next permuted watermark bit into the low frequencies of the DCT-transformed block. To increase the robustness of the embedded watermark against unintentional attacks, as explained in Section II, the locations for embedding the permuted watermark bit are restricted to the low frequency locations 0, 1, 2, and 3 as shown in Figure 1. The embedding location with the largest DCT coefficient is selected as the embedding location for a permuted watermark bit with value 1, while the embedding location with the smallest DCT coefficient is selected as the embedding location for a

permuted watermark bit with value 0. It should be noted that the largest and the smallest values of the DCT coefficients are relative values that are unlikely to change under attacks. This technique, thus, attempts to utilize the more robust relative values of the DCT coefficients instead of the absolute values that can easily change under attacks. It should be also noted that we do not need to physically embed the permuted watermark bit. We only need to remember its embedding location. This implies that in the proposed technique the embedded watermark is 100% transparent. Since there are only 4 possible embedding locations, we need only two bits to represent each embedding location. For example, (0, 0), (0, 1), (1, 0), and (1, 1) can be used to represent embedding locations 0, 1, 2, and 3 respectively. Thus, the permuted watermark bit is replaced by two bits representing its embedding location. Steps 4 and 5 are repeated until all the permuted watermark bits have been embedded in the DCT-transformed host image blocks. Since each permuted watermark bit is represented by its embedding location and since an embedding location is represented by two bits, the final result is an array with double the size of the watermark array.

- **Step 6:** After embedding all the permuted watermark bits into the DCT-transformed host image blocks, the resultant array of the embedding locations of the permuted watermark bits is decomposed into two arrays with the same size. The first array stores the first bit representing each permuted watermark bit embedding location, while the second array stores the second bit. The two arrays are of the same size as the watermark array and are considered to be the watermark shares or transparencies [17]. For example, suppose that the embedding location of the permuted watermark bit (a, b) is location 1 as shown in Figure 1. This embedding location is represented by (0, 1). Thus, the corresponding value of the bit (a, b) in share or transparency 1 is 0 and in share or transparency 2 is 1. One of the generated transparencies or shares is public, while the other is secret and is registered to the Certified Authority (CA) for additional security and protection against intentional attacks (that attempt to extract the embedded watermark for subsequent destruction) as mentioned before.
- **Step 7:** Finally, inverse DCT-transform each watermarked block from the frequency domain to the spatial domain forming the watermarked host image using the following equation [6]:

$$pixel(x, y) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} C(i) * C(j) * DCT(i, j) * \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right] \quad (3)$$

$$\text{where } C(i), C(j) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } i, j=0 \\ \sqrt{\frac{2}{N}} & \text{otherwise} \end{cases}$$

It is worth noting that in the proposed technique, this step does not need to be explicitly performed. This is because the watermark bits are not physically embedded into the host image blocks as explained above.

B. The Watermark Extraction Process

The steps of the proposed watermark extraction process are the reverse of the steps of the watermark embedding process. They can be described as follows:

- **Step 1:** Divide the watermarked host image into non-overlapping $4 * 4$ blocks.
- **Step 2:** Reassemble the two transparencies or shares to generate the permuted watermark bits embedding locations. Remember that the two values in each two corresponding locations in the two transparencies are the two values representing the embedding location of the corresponding bit of the permuted watermark. For example, suppose that the value of bit (a, b) in share 1 is 0 and that of bit (a, b) in share 2 is 1. This implies that the embedding location of bit (a, b) of the permuted watermark is that represented by (0, 1), which is location 1 as shown in Figure 1.
- **Step 3:** Arrange the watermarked host image blocks sequentially one by one and row by row. Use a pseudo-random number generator (with the same seed as the embedding pseudo-random number generator) to determine the sequence of host image blocks used for embedding (and thus extracting) the permuted watermark bits.
- **Step 4:** DCT-transform the next watermarked host image block from the spatial domain to the frequency domain.
- **Step 5:** Extract the next permuted watermark bit from the corresponding embedding location in the low frequencies of the DCT-transformed block. The corresponding bits in the reassembled transparencies are used for specifying the embedding location in the DCT-transformed block as explained above. If the embedding location has the largest value among the four low frequency locations shown in Figure 1, the embedded permuted watermark bit is a 1. Otherwise, it is a 0. Steps 4 and 5 are repeated until all the embedded permuted watermark bits have been extracted.
- **Step 6:** After extracting all the permuted watermark bits, use Torus Automorphism (TA) permutation to rearrange the watermark bits (using the same keys used for disarranging the watermark bits before embedding).

C. The Secret Keys

Several kinds of secret keys are utilized in the proposed technique. These secret keys are necessary in the proposed watermark embedding and extracting processes. They should be preserved well for watermark extraction and ownership verification and to increase the robustness of the embedded watermark against intentional attacks (that attempt to extract the embedded watermark for subsequent destruction). These secret keys are as follows:

- The two shares or transparencies needed for reassembling the embedding locations of the permuted watermark bits. The size of the transparencies is the same as that of the watermark.
- The seed of the pseudo-random number generator needed for determining the sequence of host image blocks used for embedding (and thus extracting) the permuted watermark bits.
- The k and m parameters of Torus Automorphism (TA) permutation. These are needed to disarrange the watermark bits before embedding and thus to rearrange the extracted permuted watermark bits.

It is clear that the proposed technique is blind. This is because the watermark extraction process requires only the secret key(s) for watermark embedding in addition to the two transparencies representing the embedding locations of the permuted watermark bits. Neither the original host image nor the embedded watermark bit sequence is needed.

IV. EXPERIMENTAL RESULTS

This section provides experimental results that demonstrate the robustness of the proposed technique against unintentional attacks. Two host images have been used in the experiments. These are the boat image and the baboon image shown in Figures 2 and 3 respectively. Each of these images is of size $256*256$. The rose watermark shown in Figure 4 is of size $128*128$. It has been used for watermarking these two host images.

Figures 5 and 6 show the extracted watermarks from the boat image and the baboon image respectively under different types of unintentional attacks: median filtering, blurring, sharpening, Gaussian noise addition, salt and pepper noise addition, and JPEG compression.

It is clear that the watermarks, embedded using the proposed technique, are robust to these different types of attacks. This is not only due to embedding the watermark bits in the low frequencies of the DCT-transformed blocks of the host image, but also due to intelligently utilizing the more robust relative values of the low frequency components instead of the absolute values as explained in Section III.



Figure 2. The boat image.



Figure 3. The baboon image.



Figure 4. The rose watermark.

V. DISCUSSION AND CONCLUSIONS

This paper is concerned with digital image watermarking for ownership assertion and copyright protection. A novel technique to achieve high degree of robustness with blindness in addition to transparency is proposed in this paper. The proposed technique incorporates three techniques to achieve robustness against intentional attacks: Torus Automorphism (TA) permutation to disarrange the watermark bits before embedding into the host image), a pseudo-random number generator to select a random sequence of host image blocks for embedding the scrambled watermark bits, and visual cryptography to produce two shares and register one to the Certified Authority (CA). On the other hand, to achieve robustness against unintentional attacks, the proposed technique not only watermarks the low frequency DCT components of the host image blocks, but also intelligently utilizes the more robust relative values of these low frequencies instead of their absolute values.

To achieve blindness, the technique is designed such that it only requires few keys and two transparencies for watermark extraction. In the literature, the size of the shares is usually double the size of the host image and consequently the process of generating the shares is usually time-consuming. In the proposed technique, generating the shares is straightforward and fast and the size of each share is exactly equal to the watermark size.

In addition to the above advantages, the proposed technique is designed to be 100% transparent so as not to degrade the quality of the host image. Embedding the watermark relies on

remembering the embedding locations within the low frequency coefficients and not on changing the pixels values of these coefficients.

The robustness of the proposed technique against unintentional attacks is demonstrated using two test images under several attacks: median filtering, blurring, sharpening, Gaussian noise addition, salt and pepper noise addition, and JPEG compression with different quality factors 30, 50, and 70.

In summary, the paper has the following contributions:

- Combining the contradictory goals of robustness against intentional and unintentional attacks with transparency in addition to blindness.
- Achieving 100% transparency of the embedded watermark.
- Achieving high degree of robustness against several types of unintentional attacks.
- The technique is blind and requires only 3 keys and two transparencies.
- Using three techniques to achieve high robustness against intentional attacks.
- Unlike other techniques in the literature, generating the transparencies is straightforward and fast.
- Unlike other techniques in the literature where the size of each share is usually equal double the size of the host image, the size of each share in the proposed technique is exactly equal to the watermark size

REFERENCES

- [1] C.Y. Lin and S.F. Chang, "A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 2, pp. 153-168, 2001.
- [2] C. Yanga, W. Hua, and C. Lin, "Reversible Data Hiding by Coefficient-Bias Algorithm," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 2, pp. 91-100, 2010.
- [3] C. Chang and J. Chuang, "An Image Intellectual Property Protection Scheme for Gray-Level Images Using Visual Secret Sharing Strategy," *Pattern Recognition Letters*, vol. 23, pp. 931-941, 2002.
- [4] M. Hua, D. Loub, and M. Chang, "Dual-Wrapped Digital Watermarking Scheme for Image Copyright Protection," *Computers & Security*, vol. 26, pp. 319-330, 2007.
- [5] J. Hussein, "Spatial Domain Watermarking Scheme for Colored Images Based on Log-Average Luminance," *Journal of Computing*, vol. 2, no. 1, 2010.
- [6] S. Saryazdi and M. Demehri, "A Blind DCT Domain Digital Watermarking," in *proceedings of the 3rd International Conference on Sciences of Electronic, Technologies of Information and Telecommunications*, 2005.
- [7] S. D. Lin, S. Shie, and J.Y. Guo, "Improving the Robustness of DCT-Based Image Watermarking against JPEG Compression," *Computer Standards & Interfaces*, vol. 32, pp. 54-60, 2010.
- [8] J. Patra, J. E. Phua, and C. Bornand, "A Novel DCT Domain CRT-Based Watermarking Scheme for Image Authentication Surviving JPEG Compression," *Digital Signal Processing*, 2010.
- [9] C. Temi, S. Choomchuay, and A. Lasakul, "A Robust Image Watermarking Using Multiresolution Analysis of Wavelet," in *proceedings of ISCIT2005*, 2005.

- [10] C. Yongqiang, Z. Yanqing, and P. Lihua, "A DWT Domain Image Watermarking Scheme Using Genetic Algorithm and Synergetic Neural Network," in proceedings of the 2009 International Symposium on Information Processing (ISIP'09), pp. 298-301, 2009.
- [11] E. Ganic and A. M. Eskicioglu, "Robust DWT-SVD Domain Image Watermarking: Embedding Data in All Frequencies," in proceedings of the 2004 workshop on Multimedia and security, 2004.
- [12] S. Bedi, A. Kumar, and P. Kapoor, "Robust Secure SVD Based DCT-DWT Oriented Watermarking Technique for Image Authentication," in proceedings of International Conference on IT to Celebrate S. Charmonman's 72nd Birthday, 2009.
- [13] D. Lou, H. Tso, and J. Liu, "A Copyright Protection Scheme for Digital Images Using Visual Cryptography Technique," Computer Standards & Interfaces, vol. 29, pp. 125-131, 2007.
- [14] G. Voyatzis and I. Pitas, "Applications of Toral Automorphisms in Image Watermarking," In proceedings of the Image Processing International Conference, 1996.
- [15] C. Chang, J.Y. Hsiao, and C.L. Chiang, "An Image Copyright Protection Scheme Based on Torus Automorphism," in proceedings of the 1st International Symposium on Cyber Worlds, pp. 217-224, 2002.
- [16] M. Engedy, V.N.K. Munaga, and A. Saxena, "A Robust Wavelet Based Digital Watermarking Scheme Using Chaotic Mixing," in proceedings of the 1st International Conference on Digital Information Management, pp. 36-40, 2006.
- [17] N. Naor and A. Shamir, "Visual Cryptography," in Proceedings of Eurocrypt'94, pp. 1-12, 1994.
- [18] C. Wang, S. Tai, and C. Yu, "Repeating Image Watermarking Technique by the Visual Cryptography," IEICE Transactions on Fundamentals, pp. 1589-1598, 2000.

AUTHORS PROFILE

Hanan Elazhary received her B.Sc. degree in Electronics and Communications Engineering and her M.Sc. degree in Computer Engineering from the faculty of Engineering, Cairo University, Egypt. She received her Ph.D. degree in Computer Science and Engineering from the University of Connecticut, USA. Currently, she is working as a researcher at the Electronics Research Institute, Cairo, Egypt. She is also working as an assistant professor at Akhbar Elyom Academy, Cairo, Egypt. Her research interests include high performance computing (HPC), software engineering, artificial intelligence, and computer networks. Elazhary has supervised several graduation projects and is currently supervising one Ph.D. student in the field of high-performance computing.

Sawsan Morkos received her B.Sc. degree in Electronics and Communications Engineering and her M.Sc. degree and Ph.D. degree in Computer Engineering from the faculty of Engineering, Cairo University, Egypt. Currently, she is working as a researcher at the Electronics Research Institute, Cairo, Egypt. Her research interests include image processing and artificial intelligence.

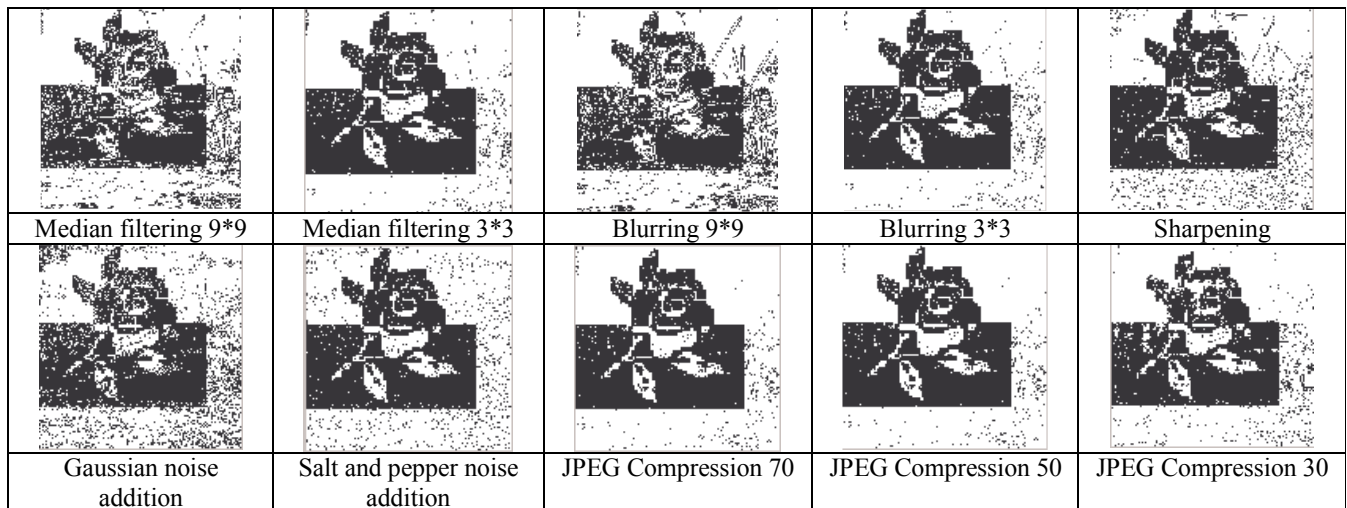


Figure 5. The extracted watermarks from the boat image under different attacks.

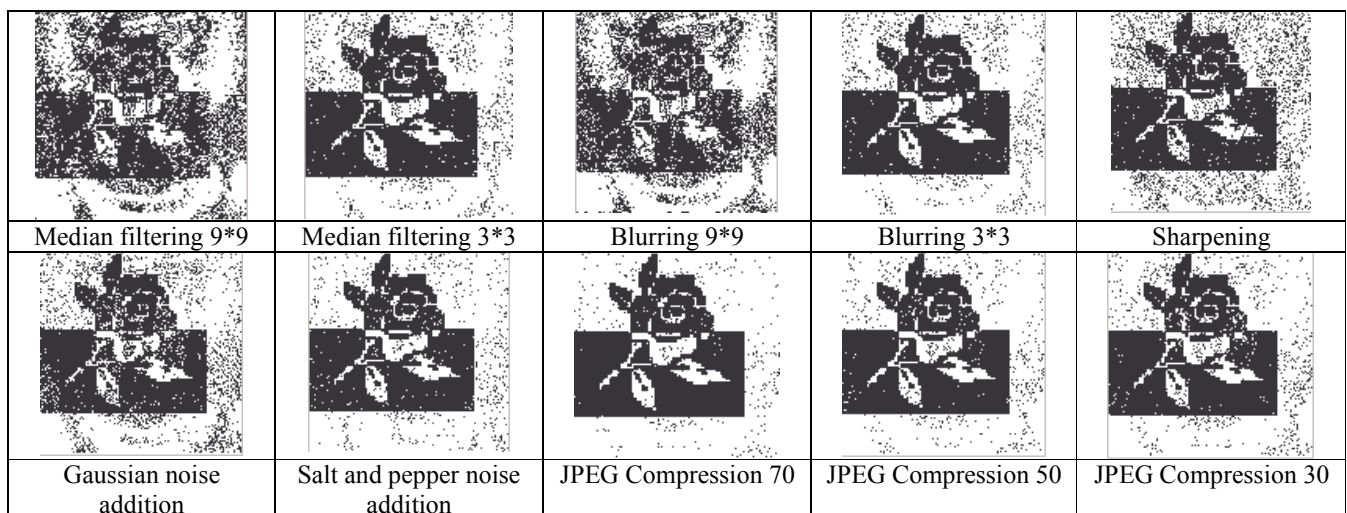


Figure 6. The extracted watermarks from the baboon image under different attacks.

An Enhanced LEACH Protocol using Fuzzy Logic for Wireless Sensor Networks

J.Rathi/Lecturer
Dept.of.B.Sc(CT)
K.S.Rangasamy college of technology
Tiruchengode-637215
Tamilnadu, India.

Dr.G.Rajendran/ Prof and Head
Dept.of MATHS
Kongu Engg. College
Perundurai, Erode-638 052,
Tamilnadu, India.

Abstract--The Wireless Sensor Networks consists of a large number of small and cheap sensor nodes that have very restricted energy, processing power and storage. They usually examine areas, collect data and report to the base station (BS). Due to the achievement in low-power digital circuit and wireless communication, many applications of the WSN are developed and already been used in habitat monitoring, military object and object tracking. The disadvantage in this monitoring leads to clustering the networks. The hierarchical network structures which are created by clustering technique are called clusters. Clusterhead is elected by its nearest networks. Clusterhead selection becomes a significant problem because of its dynamic environment. In this paper, the problem of suitable clusterhead selection in wireless sensor networks is analyzed. Appropriate cluster-head node election can significantly reduce the energy consumption and enhance the lifetime of the network. The fuzzy logic technique for clusterhead selection is proposed in this paper based on three descriptors, namely, Energy, Concentration and Density. The experimental results shows the substantial increase in the network lifetime depends on network configuration as compared to probabilistically selecting the nodes as cluster-heads using only local information.

Keywords— Wireless Sensor Networks, Fuzzy Logic, sensor networks, Cluster head

I. INTRODUCTION

WIRELESS sensor networks (WSN) are composed of a compilation of devices that communicate with each other over a wireless medium. Such a kind of sensor network forms spontaneously whenever devices are in transmission range. Joining and leaving of nodes occurs dynamically, particularly when they are like mobile devices. Potential applications of wireless sensor networks can be found in traffic scenarios, ubiquitous Internet access, collaborative work, and many more. Wireless sensor networks assemble and process environmental data. They consist of small devices communicating through radio. Normally, data processing in Wireless Sensor Networks occurs locally and decentralized. The architecture of the model is shown in Figure 1.

In wireless sensor networks [5] [7], clustering is one of the mainly popular techniques for locality-preserving network organization. Cluster-based architectures

effectively decrease energy consumption, and enable efficient realization and routing protocols, data aggregation, and security mechanisms.

A cluster [6] [17] is a collection of interconnected nodes with a dedicated node called clusterhead. Clusterheads are accountable for cluster management, such as scheduling of the medium access, dissemination of control messages, or data aggregation. Therefore, the role of the clusterhead is critical for the appropriate network operation. Failure of a clusterhead leads to expensive clusterhead re-election and re-clustering operations.

In stagnant networks, the role of the clusterhead may be assigned to any node in the cluster in a self-organized way. Often, this role is assigned in turn to the nodes in order to ensure fairness, as a clusterhead consumes more energy than a regular sensor node. An essential criterion for the clusterhead selection is the remaining energy level of the node. However, for fault-tolerant clusterhead selection in dynamic networks, some additional criteria for choosing a clusterhead are required. For example, considering node mobility, if a clusterhead is close to the network partition border, it may disappear from the cluster earlier than a more centrally located node. On the other hand, a centrally located node should not be selected as a clusterhead if its failure leads to cluster partitioning.

The energy utilization can be minimized by allowing only a portion of the nodes, which called cluster heads, to communicate with the base station. The data sent by each node is then composed by cluster heads and compressed. After that the aggregated data is transmitted to the base station. Although clustering can reduce energy consumption [8] [9], it has certain limitations. The main setback is that energy consumption is concentrated on the cluster heads [4]. In order to overcome this demerit, the issue in cluster routing of how to distribute the energy consumption [10] must be resolved. The representative solution is LEACH (Low Energy Adaptive Clustering Hierarchy), which is a localized clustering method based on a probability

model. The main idea of LEACH procedure is that all nodes are chosen to be the cluster heads periodically, and each period contains two stages:

- Construction of clusters
- Data communication

Cluster heads are selected according to the probability of optimal cluster heads determined by the networks. After the selection of cluster heads, the clusters are constructed and the cluster heads communicate data with base station. Because LEACH is only depend on probability model, some cluster heads may be very close to each other and can be located in the edge of the WSN.

These disorganized cluster heads could not maximize energy efficiency. To overcome the defects of LEACH methodology, a cluster head election method using fuzzy logic has been introduced. This method proved that the network lifetime can be efficiently prolonged by using fuzzy variables (concentration, energy and density). In this proposed method, a part of energy is spent to get the data of the three variables especially concentration and density. The experimental show that the proposed approach increases the network lifetime significantly when compared to LEACH approach.

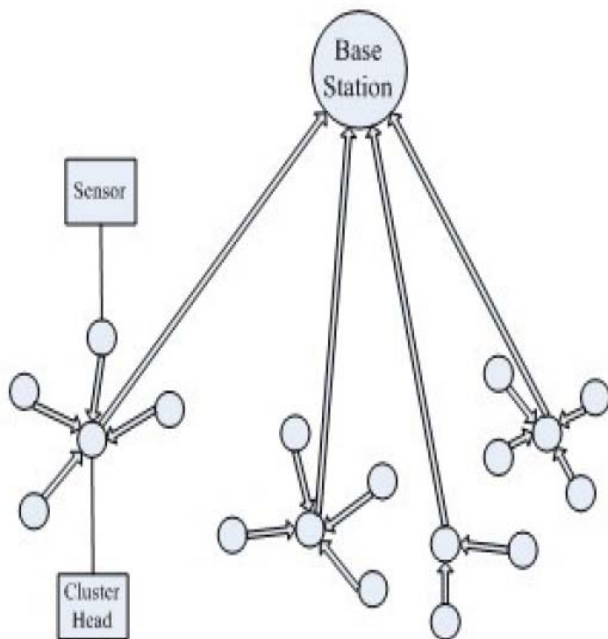


Fig. 1: WSN Architecture

In this paper, a method based on LEACH using Fuzzy Logic to cluster heads selection is proposed based on three variables - battery level of node, node density and distance from base station, and this method will be introduced based on the assumption that the WSN can get their coordinate. Although this method has the same drawback as of Gupta's method, it presents a better

result. For a cluster, the nodes selected by the base station are the nodes that have the higher chance to become the cluster heads using Fuzzy Logic based on their battery level, node density and distance.

II. RELATED WORKS

Handy *et al.*, [1] proposed the Low Energy Adaptive Clustering Hierarchy with Deterministic Cluster-Head Selection. This paper focuses on reducing the power [11] [13] consumption of wireless microsensor networks. Therefore, a communication protocol named LEACH (Low-Energy Adaptive Clustering Hierarchy) is modified. The author extend LEACH's stochastic clusterhead selection algorithm by a deterministic component. Depending on the network configuration an increase of network lifetime by about 30 % can be accomplished. Furthermore, a new approach is presented to define lifetime of microsensor networks using three new metrics FND (First Node Dies), HNA (Half of the Nodes Alive), and LND (Last Node Dies).

W. Heinzelman *et al.*, [2] presented an Energy-efficient Communication Protocol for Wireless Microsensor Networks. In this paper, the author looks at communication protocols, which can have significant impact on the overall energy dissipation of these networks. Based on the findings that the conventional protocols of direct transmission, minimum-transmission-energy, multihop routing, and static clustering may not be optimal for sensor networks, the author propose LEACH (Low-Energy Adaptive Clustering Hierarchy), a clustering-based protocol that utilizes randomized rotation of local cluster base stations (cluster-heads) to evenly distribute the energy load among the sensors in the network. LEACH uses localized coordination to enable scalability and robustness for dynamic networks, and incorporates data fusion into the routing protocol to reduce the amount of information that must be transmitted to the base station. Simulations show that LEACH can achieve as much as a factor of 8 reductions in energy dissipation compared with conventional routing protocols. In addition, LEACH is able to distribute energy dissipation evenly throughout the sensors, doubling the useful system lifetime for the networks we simulated.

Shen *et al.*, [3] suggested the Sensor Information Networking Architecture and applications; this paper introduces a sensor information networking architecture, called SINA that facilitates querying, monitoring, and tasking of sensor networks. SINA serves the role of middleware that abstracts a network of sensor nodes as a collection of massively distributed objects. SINA's execution environment provides a set of configuration

and communication primitives that enable scalable and energy-efficient organization of and interactions among sensor objects. On top the execution environment is a programmable substrate that provides mechanisms to create associations and coordinate activities among sensor nodes. Users then access information within a sensor network using declarative queries, or perform tasks using programming script

III. METHODOLOGY

In this paper the cluster-heads are elected by the base station in each round by calculating the chance each node has to become the cluster-head by considering three fuzzy descriptors:

- Node concentration
- Energy level in each node
- Node Density

In the proposed approach, the better cluster-heads are produced by the central control algorithm in the base station. This is because the global knowledge about the network is contained in base station. In addition, base stations are many times more potent than the sensor nodes, having sufficient memory, power and storage. In the proposed approach energy is spent to transmit the location information of all the nodes to the base station. Considering WSNs are meant to be deployed over a geographical area with the main purpose of sensing and gathering information, this paper assumes that nodes have minimal mobility, thus sending the location information during the initial setup phase is sufficient.

The cluster-head collects n number of k bit messages from n nodes that joins it and compresses it to cn k bit messages with $c \leq 1$ as the compression coefficient. The operation of this fuzzy cluster-head election scheme is divided into two rounds each consisting of a setup and steady state phase similar to LEACH. During the setup phase the cluster-heads are determined by using fuzzy [14] knowledge processing and then the cluster is organized. In the steady state phase the cluster-heads collect the aggregated data and performs signal processing functions to compress the data into a single signal. This composite signal is then sent to the base station.

The radio model used here is with $E_{elec} = 50$ nJ/bit as the energy dissipated by the radio to run the transmitter or receiver circuitry and $\epsilon_{amp} = 100$ pJ/bit/m² as the energy dissipation of the transmission amplifier.

The energy expended during transmission and reception for a k bit message to a distance d between transmitter and receiver node is given by:

$$E_{Tx}(k, d) = E_{elec} * k + \epsilon_{amp} * k * d^\alpha$$

$$E_{Rx}(k) = E_{elec} * k$$

Where, λ is the path loss exponent and $\lambda \geq 2$.

The model of fuzzy logic control consists of a fuzzifier, fuzzy rules, fuzzy inference engine, and a defuzzifier. The most commonly used fuzzy inference technique called Mamdani Method is used in the proposed approach due to its simplicity. The process is performed in four steps:

- **Fuzzification** of the input variables energy, concentration and density - taking the crisp inputs from each of these and determining the degree to which these inputs belong to each of the appropriate fuzzy sets.
- **Rule evaluation** - taking the fuzzified inputs, and applying them to the antecedents of the fuzzy rules. It is then applied to the consequent membership function (Table 1).
- **Aggregation of the rule outputs** - the process of unification of the outputs of all rules.
- **Defuzzification** - the input for the defuzzification process is the aggregate output fuzzy set chance and the output is a single crisp number.

During defuzzification, it finds the point where a vertical line would slice the aggregate set chance into two equal masses. In practice, the COG (Center of Gravity) is calculated and estimated over a sample of points on the aggregate output membership function, using the following formula:

$$COG = (\sum \mu_A(x) * x) / \sum \mu_A(x)$$

Where, $\mu_A(x)$ is the membership function of set A.

Expert knowledge is represented based on the following three descriptors:

- **Node Energy** - energy level available in each node, designated by the fuzzy variable energy,
- **Node Concentration** - number of nodes present in the vicinity, designated by the fuzzy variable concentration,
- **Node Density** - density of node in the cluster

The linguistic variables used to represent the node energy and node concentration, are divided into three levels: low, medium and high, respectively, and there are three levels to represent the node density: sparse, medium and dense respectively. The outcome to represent the node cluster-head election chance was divided into seven levels: very small, small, rather small, medium, rather large, large, and very large. The fuzzy rule base currently includes rules like the following: if the energy is high and the concentration is high and the density is close then the node's cluster-head election chance is very large.

Thus, $3^3 = 27$ rules are used for the fuzzy rule base. In this paper, the triangle membership functions are used to

represent the fuzzy sets medium and adequate and trapezoid membership functions to represent low, high, close and far fuzzy sets. The membership functions developed and their corresponding linguistic states are represented in Table 1 and Figures 2 through 5.

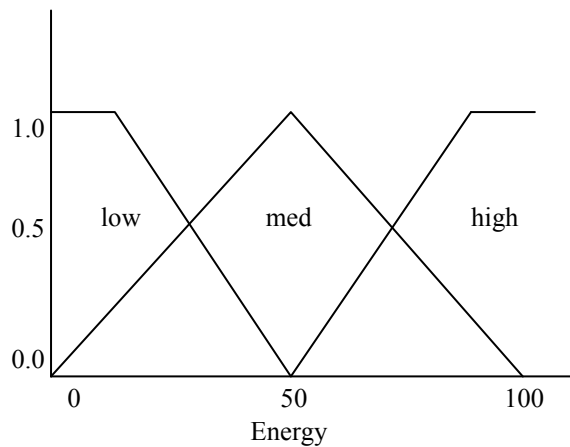


Figure2. Fuzzy set for fuzzy variable energy

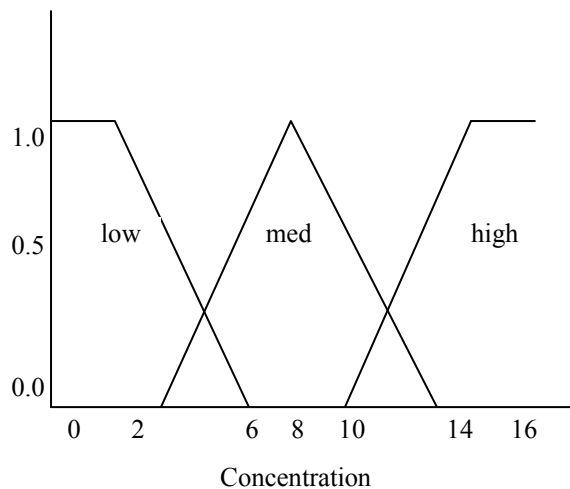


Figure3. Fuzzy set for fuzzy variable concentration

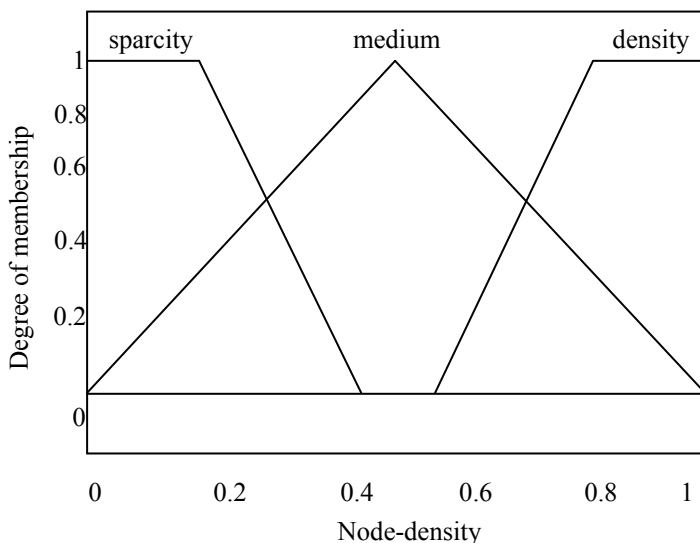


Figure4. Fuzzy set for fuzzy variable density

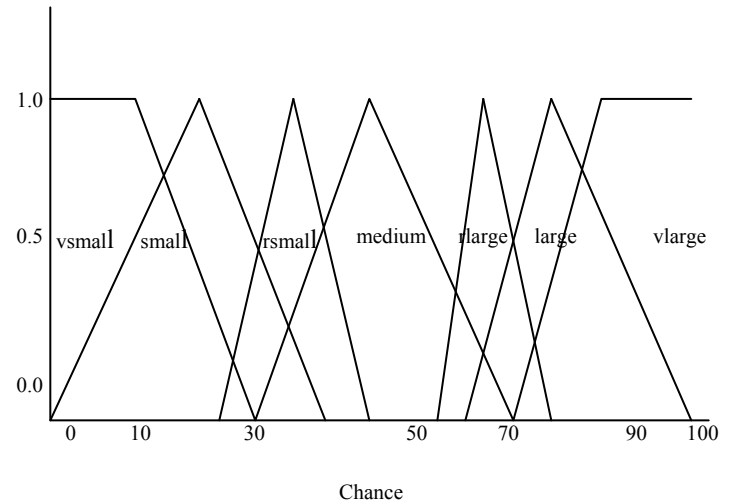


Figure5. Fuzzy set for fuzzy variable chance

TABLE1: FUZZY RULE BASE

S.no	Energy	Concentration	Density	Chance
1	low	low	sparse	small
2	low	low	medium	small
3	low	low	dense	vsmall
4	low	med	sparse	small
5	low	med	medium	small
6	low	med	dense	small
7	low	high	sparse	rsmall
8	low	high	medium	small
9	low	high	dense	vsmall
10	med	low	sparse	rlarge
11	med	low	medium	med
12	med	low	dense	small
13	med	med	sparse	large
14	med	med	medium	med
15	med	med	dense	rsmall
16	med	high	sparse	large
17	med	high	medium	rlarge
18	med	high	dense	rsmall
19	high	low	close	rlarge
20	high	low	adeq	med
21	high	low	far	rsmall
22	high	med	close	large
23	high	med	adeq	rlarge
24	high	med	far	med
25	high	high	close	vlarge
26	high	high	adeq	rlarge
27	high	high	far	med

Legend: med-medium, vsmall-very small, rsmall-rather small, vlarge-very large, rlarge-rather large.

All the nodes are compared on the basis of chances and the node with the maximum chance is then elected as the cluster-head. Each node in the cluster associates itself to the cluster-head and starts transmitting data. The

data transmission phase is similar to the LEACH steady-state phase.

IV. EXPERIMENTAL RESULTS

The different experimental are conducted on the proposed system and the results are discussed in this section.

The reference network consists of 100 nodes randomly distributed over an area of 400X400 meters. The base station is located at 200, 450. In the first phase of the simulation each node has a random energy between 0 and 100. The base station computes the concentration for each node by calculating the number of other nodes within the area of 20X20 meters by considering its density as well. The values are then fuzzified and passed to the fuzzy rule base for rule evaluation. After this, defuzzification gives the cluster-head election chance. If the chance is large or very large, then that node is chosen as a cluster center. This techniques shows better finding of cluster than the conventional methods such as LEACH, etc,

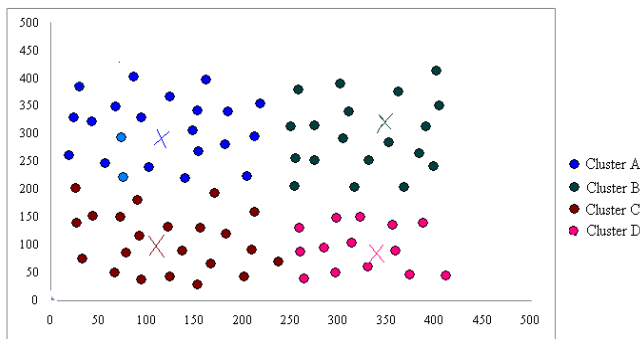
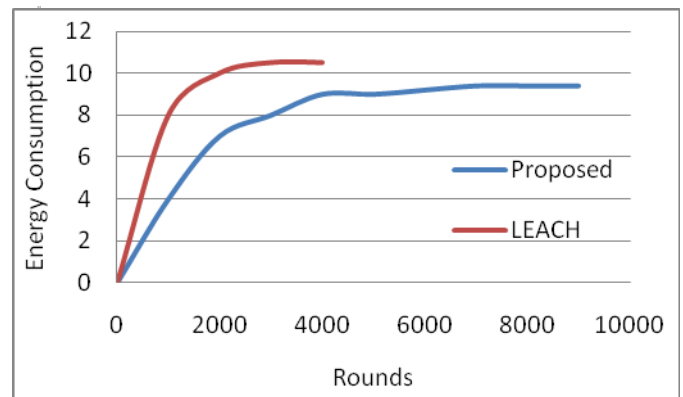


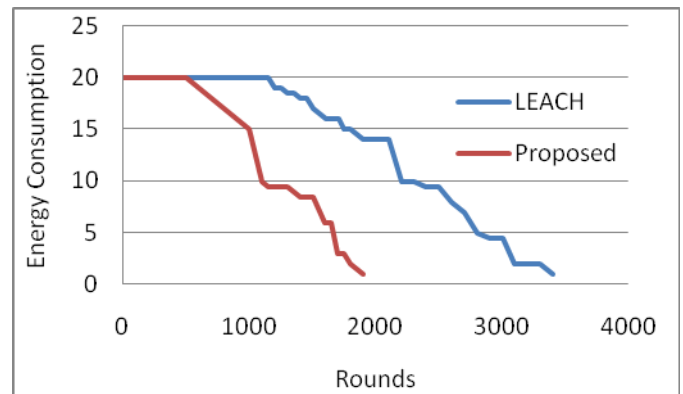
Fig 6: Cluster formation of the simulated network using 4 clusters and a network size of 400x400 meters.

Figure 7(a) shows the energy consumption of the proposed system compared with that of LEACH. The increase rate of energy consumption of the proposed system is much lower than the rate of LEACH. When LEACH has used all of its energy and demises, the proposed approach still has 54% of its energy left.

Figure 7(b) shows the nodes alive of the proposed system compared with the nodes alive of LEACH. Besides, both the dead time of the first node and the dead time of the last node of proposed system are later than those of LEACH. Thus it is clear that, compared to LEACH; the proposed has approximately 88% of its nodes alive. So WSN can get longer life and enjoy longer receiving of integral data by using the proposed method.



(a) Energy Consumption



(b) Nodes

Fig. 7: (a) Energy Consumption (b) Nodes

This results in a situation where the BS can receive at least 9000 more messages from the network before all energy is consumed. The energy consumed in the network is evenly distributed among the nodes in AROS. Clusters far away from the BS in the proposed system will survive until the end and continue to gather information.

V. CONCLUSION

The new approach for cluster-head election for Wireless Sensor Networks (WSN) is presented in this paper. Cluster-heads were elected by the base station in each round by calculating the chance each node has to become the cluster-head using three fuzzy descriptors: node energy, node concentration and node density. The energy is the most important factor in designing the protocol for WSN. The propose approach achieved better reduction in the usage of energy for finding center of cluster. The simulation result shows that the proposed approach has good energy consumption when compared to LEACH methodology. By the proposed method, the

better network life time is accomplished when compared to LEACH.

REFERENCES

- [1] M.J. Handy, M. Haase and D. Timmermann, "Low energy adaptive clustering hierarchy with deterministic cluster-head selection," in Proc. 4th International Workshop on Mobile and Wireless Communications Network, Sept. 2002, pp. 368 - 372.
- [2] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-efficient Communication Protocol for Wireless Microsensor Networks", Proceedings of 33rd Hawaii International Conference on System Sciences, Jan., 2000
- [3] C. Shen, C. Srisathapornphat, and C. Jaikaeo, Sensor Information Networking Architecture and applications, IEEE Personal Communications, pp. 52-59, Aug 2001
- [4] J. Anno, L. Barolli, F. Xhafa, and A. Durresi, A Cluster Head Selection Method for Wireless Sensor Networks Based on Fuzzy logic, Proc. of IEEE TENCON-2007, CD-ROM, 4 pages, 2007
- [5] C. Chee-Yee and S.P. Kumar, "Sensor networks: evolution, opportunities, and challenges," in Proc of the IEEE, Aug. 2003, pp.1247 - 1256.
- [6] Q. Liang, "Clusterhead election for mobile ad hoc wireless network," in Proc. 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC), Sept. 2003, pp. 1623 - 1628.
- [7] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "An application-specific protocol architecture for wireless micro sensor networks," in IEEE Transactions on Wireless Communications, Oct. 2002, pp. 660 - 670.
- [8] S. Lindsey, C. Raghavendra, K. Sivalingam, "Data gathering in sensor networks using the energy delay metric", Proc. of the IPDPS Workshop on Issues in Wireless Networks and Mobile Computing, April 2001.
- [9] V. Rodoplu, T. H. Meng, "Minimum energy mobile wireless networks", IEEE Jour. Selected Areas Comm., August 1999, pp. 1333-1344.
- [10] L. Zhong, R. Shah, C. Guo, J. Rabaey, "An ultra low power and distributed access protocol for broadband wireless sensor networks", IEEE Broadband Wireless Summit, Las Vegas, May 2001.
- [11] J. M. Rabaey, M. J. Ammer, J. L. da Silva, D. Patel, S. Roundy, "PicoRadio supports ad hoc ultra-low power wireless networking, IEEE Computer, July 2000, pp. 42-48.
- [12] K. Pister, "On the limits and applications of MEMS sensor networks", Defense Science Study Group report, Institute for Defense Analysis, Alexandria, VA.
- [13] R. Min, M. Bhardwaj, S. Cho, E. Shih, A. Sinha, A. Wang, A. Chandrakasan, "Low-power wireless sensor networks", VLSI Design 2001, Invited Paper, Bangalore, January 2001.
- [14] I. Gupta, D. Riordan, and S. Sampalli, "Cluster-head election using fuzzy logic for wireless sensor networks," Communication Networks and Services Research Conference, 2005. Proceedings of the 3rd Annual, pp. 255-260, 2005.
- [15] O. Dousse, P. Thiran, and M. Hasler, "Connectivity in Ad hoc and Hybrid Networks," presented at INFOCOM'02, 2002.
- [16] P. Santi, "Topology Control in Wireless Ad Hoc and Sensor Networks" Wiley, 2005.
- [17] K. Dasgupta, K. Kalpakis, and P. Namjoshi, "An efficient clustering based heuristic for data gathering and aggregation in sensor networks," Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE, vol. 3, pp. 1948-1953, 2003.

A Novel Approach for Hiding Text Using Image Steganography

Sukhpreet Kaur*

Department of Computer Science and Engineering
Baba Farid College of Engineering and Technology
Bathinda-151001, Punjab, India

Sumeet Kaur

Department of Computer Engineering
Yadavindra College of Engineering Punjabi University
Guru Kashi Campus
Talwandi Sabo, Punjab, India

Abstract— With the increasing use of internet for communication, the major concern of these days is, the security of data being communicated over it. Steganography is the art and science of invisible communication. It hides secret information in other information, thus hiding the existence of the communicated information. In this paper we have discussed a technique of hiding text messages in the images using image steganography. The technique uses matching of secret data with pixel values of cover image as base concept. The LSBs of matched pixels are changed to mark presence of data inside that pixel. For making selection of channels for marking presence of data, a pseudo random number generator is used, which adds another layer of security to the technique and makes the extraction of secret data very difficult for the intruders. The results show that technique provides more security against visual and statistical attacks and attempts to provide more data hiding capacity by using more bits per pixel.

Keywords- *Steganography; image steganography; attacks; PSNR; security*

I. INTRODUCTION

Steganography can be defined as the technique used to embed data or other secret information inside some other object commonly referred to as cover, by changing its properties. The purpose of steganography is to set up a secret communication path between two parties such that any person in the middle cannot detect its existence; the attacker should not gain any information about the embedded data by simply looking at cover file or stego file. Steganography is the art of hiding information in ways that prevent the detection of hidden messages. Steganography, derived from Greek, literally means “covered writing.” It includes a vast array of secret communications methods that conceal the message’s very existence. These methods include invisible inks, microdots, character arrangement, digital signatures, covert channels, and spread spectrum [2]. Steganography is commonly misinterpreted to be cryptography or watermarking. While they are related in many ways, there is a fundamental difference in the way they are defined and the

problems to which they are applied. Cryptography protects the secret data by making it difficult to understand by the intruder but still the intruder knows that the secret data exists, so he will try his best to decode the data. Steganography & encryption are both used to ensure data confidentiality however the main difference between them is that with encryption anybody can see that both parties are communicating in secret. Steganography hides the existence of a secret message and in the best case nobody can see that both parties are communicating in secret. Watermarking is used primarily for identification and entails embedding a unique piece of information within a medium without noticeably altering the medium. Steganography uses a basic model to hide data inside the cover objects as shown in Fig. 1.

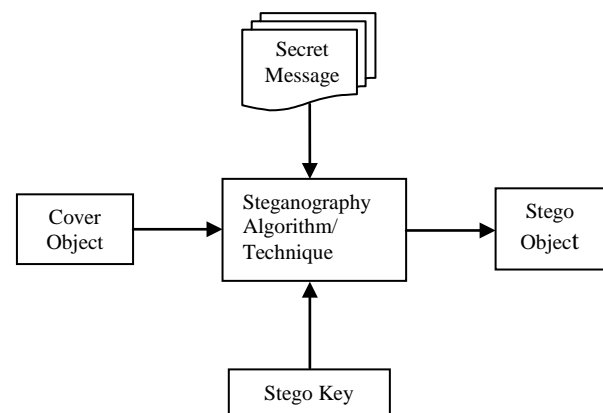


Figure 1. Basic steganography model

The basic model of steganography uses a cover object i.e. any object that can be used to hold secret information inside, the secret message i.e. the secret information that is to be sent to some remote place secretly, a stego key that is used to encode the secret message to make its detection difficult and a

steganography algorithm/technique i.e. the procedure to hide secret message inside cover object. The outcome of the process is the stego object i.e. the object that has the secret message hidden inside. This stego object is sent to the receiver where receiver will get the secret data out from the stego image by applying decoding algorithm/ technique.

In modern era, steganography is implemented by using digital media. Secret message is embedded inside digital cover media like text, images, audio, video or protocols depending upon requirement and choice of the sender. Among other types of steganography, image steganography is most widely used. The reason behind the popularity of image steganography is the large amount of redundant information present in the images that can be easily altered to hide secret messages inside them.

A. Applications of Steganography

Steganography has a wide range of applications. The major application of steganography is for secret data communication. Cryptography is also used for the same purpose but steganography is more widely used technique as it hides the existence of secret data. Another application of steganography is feature tagging. Captions, annotations, time stamps, and other descriptive elements can be embedded inside an image, such as the names of individuals in a photo or locations in a map. A secret copyright notice or watermark can be embedded inside an image to identify it as intellectual property. This is the watermarking scenario where the message is the watermark.

Steganography can be also used to combine explanatory information with an image (like doctor's notes accompanying an X-ray). Steganography is used by some modern printers, including HP and Xerox brand color laser printers. Tiny yellow dots are added to each page. The dots are barely visible and contain encoded printer serial numbers, as well as date and time stamps. The list of applications of image steganography is very long.

II. IMAGE STEGANOGRAPHY

Image steganography uses images as the cover object to hide the secret data. Images are the most widely used cover objects as they contain a lot of redundant information. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the object's use and display [3]. The redundant bits of an object are those bits that can be altered without the alteration being detected easily [5]. Image files fulfill this requirement so they are very commonly used as a medium for steganography. Audio files also contain redundant information but not used as widely as image files. A number of techniques have been proposed to use images as cover files. These techniques can be categorized in the following two ways:

- Spatial domain techniques
- Transform domain techniques

Image – also known as spatial – domain techniques embed messages in the intensity of the pixels directly, while for transform – also known as frequency – domain, images are first transformed and then the message is embedded in the image [1].

In spatial domain methods a steganographer modifies the secret data and the cover medium in the spatial domain, which involves encoding at the level of the LSBs [6]. The best widely known steganography algorithm is based on modifying the least significant bit layer of images, hence known as the LSB technique. Spatial domain algorithms embed data by substituting carefully chosen bits from the cover image pixels with secret message bits. LSB technique is the most widely used technique of image steganography. In this technique the least significant bit of all the cover image pixels is replaced with the message bits. In a 24-bit image each pixel contains 3 bytes (one for each Red, Green and Blue component), so we can store 3 bits in each pixel. Some algorithms use all pixels to hide data bits, while others use only specific areas of image. Our proposed technique is also based on the LSB method to show existence of data in a particular channel.

Transform domain techniques first transform the cover images and then hide the data inside them. Transform domain techniques [7] hide data in mathematical functions that are in compression algorithms. Discrete Cosine Transform (DCT) technique is one of the commonly used transform domain algorithm for expressing a waveform as a weighted sum of cosines. The data is hidden in the image files by altering the DCT coefficient of the image. Specifically, DCT coefficients which fall below a specific threshold are replaced with the secret bits. Taking the inverse transform will provide the stego image. The extraction process consists in retrieving those specific DCT coefficients. Jpeg Steganography is the most common example of transform domain technique of image steganography.

A good technique of image steganography aims at three aspects. First one is capacity, i.e. the maximum data that can be stored inside cover image. Second one is the imperceptibility, i.e. the visual quality of stego image after data hiding and the last is robustness i.e. security against attacks [4].

III. PROPOSED TECHNIQUE

LSB encoding is a method that claims to provide good capacity and imperceptibility. Still the existing methods do not use the full capacity of cover image. Many techniques like [8-13] have been developed to use the more and more number of bits per pixel to achieve more data hiding capacity. We have developed a technique for hiding text using image steganography that use 7 bits per pixel to hide data and still no visual changes in the stego image. We convert the messages into ASCII code and then 7 bit ASCII code of each letter is matched with pixel values of cover image. To mark the presence of data in a particular pixel we use LSB method. Which component of the pixel contains data that will be showed by using different combinations of Least Significant Bits. As we know that each pixel of the BMP image is made up of three bytes, one for Red, one for Green and one for Blue component of the pixel. Each character of the secret message

is converted into ASCII code, which is 7-bit code. So to map a character to a pixel component we need only 7 bits of the pixel. The least significant bit of every channel is free to be used as indicator to show that data is present in this channel. We will use LSB of two channels to mark presence of data in any of the three channels. The basic technique is to convert secret message into ASCII code. To decide which channels will act as indicator channels, we will use a pseudo random number. For every character of secret message, we generate a pseudo random number, depending upon the value of pseudo random number we decide that which two channels will act as indicator channels. After generating the number, convert that into binary bit sequence. Count number of 1s present in the bit sequence and number of zeros present in the bit sequence. Also calculate the parity of the pseudo random number. Now depending upon binary bit sequence of pseudo random number following three cases will be there and one case will be used to select set of indicator channels. The selection procedure is shown in table 1.

TABLE I. CRITERIA FOR SELECTION OF INDICATOR CHANNEL

Case	Indicator channel set	Order1(if parity is even)	Order2(if parity is odd)
If no. of 1s are more than number of 0s	RG	RG	GR
If no. of 0s are more than number of 1s	GB	GB	BG
If no. of 0s are equal to number of 1s	RB	RB	BR

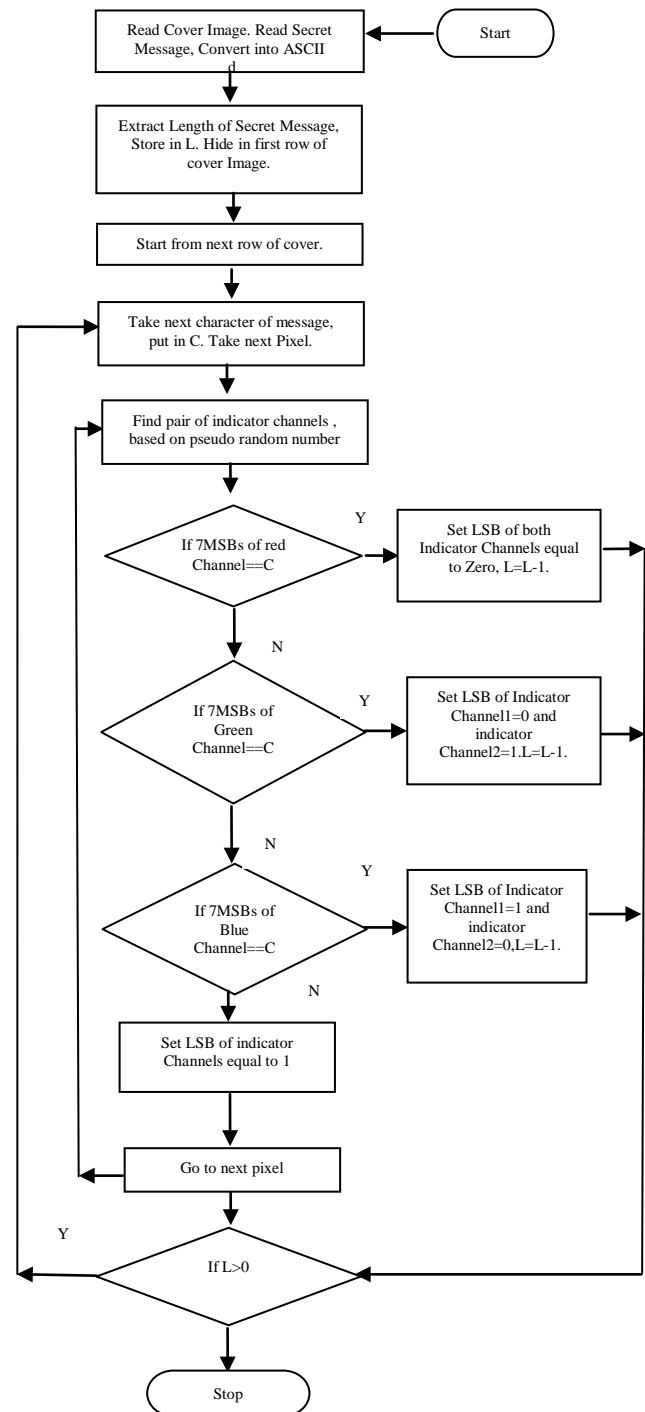
TABLE II. CRITERIA TO SET VALUE OF INDICATOR CHANNELS

Data channel (depending upon match)	LSB of indicator 1	LSB of indicator 2
RED Channel	0	0
GREEN Channel	0	1
BLUE Channel	1	0
No match	1	1

After selecting set of indicator channels we start from the first row of cover image. We hide length of secret message in first row using LSB method. Then start tracing from the second row to match first character of secret message with 7 MSBs of all three components of first pixel. If there is a match with any component then value of indicator channels is set according to

the criteria given in Table 2. As clear from the Table 2, we set the values of indicator channels as 00 if data matches in red channel, 01 if matches with green and 10 if matches with blue channel. If there is no match then the value is set as 11. Then the same procedure is repeated with the next pixel of the cover image.

A. Flow Chart of Encoding Process



B. Encoding Algorithm

The encoding part of the algorithm is as follows.

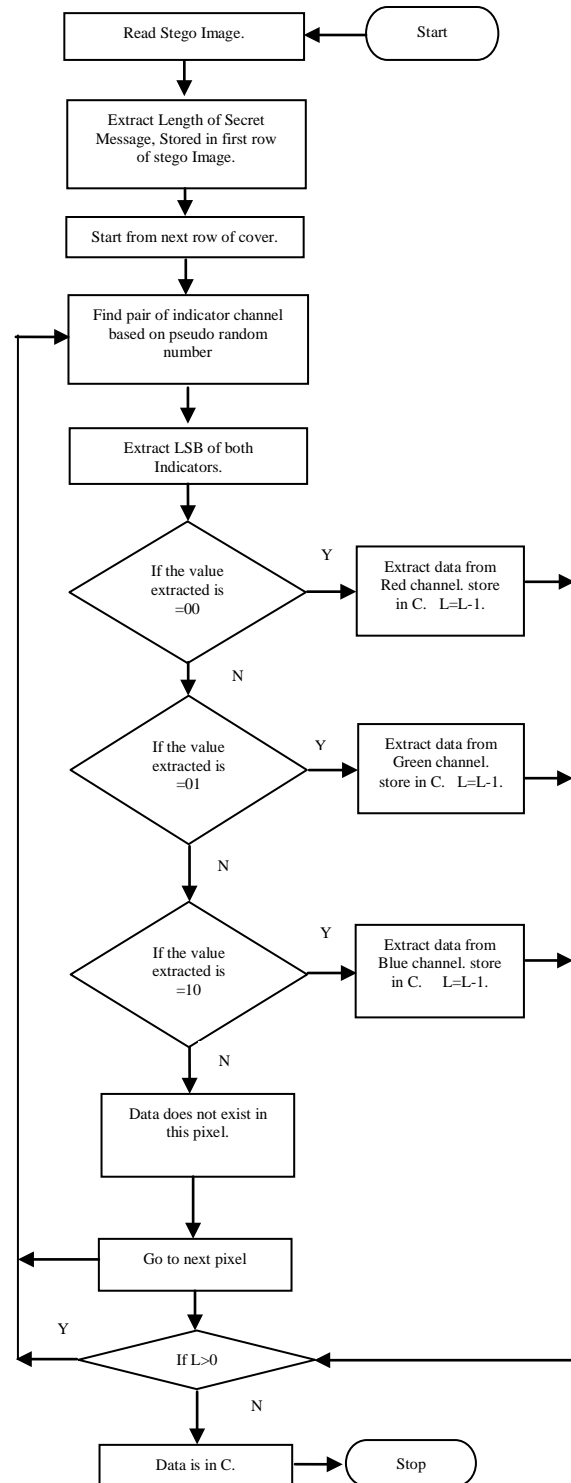
- Step 1: Read cover image.
Step 2: Read secret message, Convert into ASCII.
Step 3: Extract message length and store it into the variable L.
Step 4: Hide message length into the first row of cover image using LSB method.
Step 5: Start from second row of cover image. Take first pixel.
Step 6: Take next character from C store in a temporary variable B.
Step 7: Select indicator channel pair, depending upon the pseudo random number.
Step 8: Match the character with 7 MSBs of all three channels turn wise. If there is a match, set the value of indicator channels accordingly.
Step 9: Set $L = L - 1$. Go to step 11.
Step 10: If data not matched with any channel, set value of indicator channels equal to 1. Go to next pixel and go to step 7.
Step 11: Go to next pixel.
Step 12: Check if $L > 0$. If yes go to step no.6.
Step 13: Stop when all characters are consumed and L is equal to zero.

C. Decoding Algorithm

The decoding process will depend upon the value of the pseudo random number generator function. Number generator will generate the same numbers as it generated at sender end during decoding process. Depending upon the value of number by using table 2 we will find out the set of indicator channels. After that depending upon the value of indicator channels we will find out that data lies in which channel of which pixel. The different steps of the decoding process are as follows:

- Step 1: Read stego image.
Step 2: Read the LSB of first row to find out L.
Step 3: Start from second row of cover image. Take first pixel of second row.
Step 4: Select indicator channel pair, depending upon the pseudo random number.
Step 5: Depending upon the set of indicator channel pair, extract LSB of indicator 1 and indicator 2.
Step 6: Depending upon value of indicator channels, extract the data from pixel.
Step 7: If this value is 11 that means data does not exist in this pixel.
Step 8: Go to next pixel.
Step 9: Check if $L > 0$. If yes go to step no 4.
Step 10: Stop when all characters are retrieved and L is equal to zero.
Step 11: The values of C are in ASCII code, convert them into equivalent characters.

D. Flow Chart of Decoding Process



IV. RESULTS

To compute the performance of the proposed technique we have conducted a series of experiments. To calculate the efficiency we have used Peak Signal to Noise Ratio as major parameter. The PSNR measures Peak Signal to Noise Ratio

between two images, and higher the PSNR, the more the quality of stego image. To evaluate the results we have applied the technique to a number of colored images out of which flowers.bmp having size 32700 has been shown here to demonstrate the results achieved.



Figure 1. Cover image flowers.bmp



Figure 2. Stego image flowers.bmp

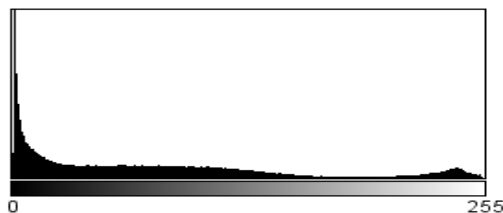


Figure 3. Histogram of cover image flowers.bmp

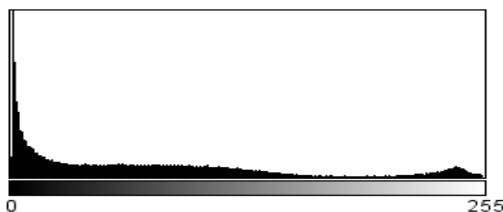


Figure 4. Histogram of stego image flowers.bmp

Fig 3. Shows the histogram of cover image and Fig 4. Shows the histogram of stego image. It is clear from the histograms that there is negligible change in the histogram of stego image. So, proposed technique is secure from statistical attacks. Table 3 shows value of PSNR after hiding messages of different sizes in the cover image flower.bmp. The results show a higher value of PSNR is achieved by the technique. Table 4 and 5 show statistical results achieved in terms of mean and standard deviation values of cover and stego images.

TABLE III. EFFECT OF INCREASE IN SIZE OF SECRET DATA ON PSNR

Image	Message Length	PSNR
flowers.bmp	100	59.638
flowers.bmp	50	61.698
flowers.bmp	25	63.372

TABLE IV. EFFECT OF INCREASE IN SIZE OF SECRET DATA ON MEAN

Image	Length of Secret Data	Mean (Cover)	Mean (Stego)	Difference
Flowers.bmp	100	75.354	75.422	0.068
Flowers.bmp	50	75.354	75.397	0.043
Flowers.bmp	25	75.354	75.384	0.030

TABLE V. EFFECT OF INCREASE IN SIZE OF SECRET DATA ON STANDARD DEVIATION

Image	Length of Secret Data	Std. Dev. (Cover)	Std. Dev. (Stego)	Difference
Flowers.bmp	100	72.174	72.159	0.015
Flowers.bmp	50	72.174	72.163	0.011
Flowers.bmp	25	72.174	72.166	0.008

V. CONCLUSION & FUTURE WORK

In this paper, we have presented a new technique to hide text inside images. The main objective was to achieve more security against statistical and visual attacks. The results show that we have been successful in achieving the same. The technique provides more security against visual attacks as the cover and stego images does not show the visible differences. The technique is also statistically secure for small text messages as there is no visible difference in the histograms of cover and stego images. We have tried to achieve more capacity by using the 7 bits per pixel to hide data. Results show a very good value of PSNR that means technique shows better imperceptibility.

The future work includes increasing the capacity further by modifying the technique. Technique can be modified to hide more data without noticeable visual changes. Some type of mapping table can be used to increase the chances of matching data with pixel values. Hence focus of the future work is to

achieve more capacity while retaining the robustness against visual attacks and statistical properties of cover image.

REFERENCES

- [1] T. Morkel, J.H.P. Eloff and M.S. Olivier, "An overview of image steganography", in Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005), Sandton, South Africa, June/July 2005.
- [2] N.F. Johnson, S. Jajodia, "Exploring steganography: seeing the unseen", Computer Journal, vol. 31, pp. 26-34, February 1998.
- [3] D.L. Currie, & C.E. Irvine, "Surmounting the effects of lossy compression on steganography", 19th National Information Systems Security Conference, pp. 194-201, 1996.
- [4] H. Zhang and H. Tang, "A novel image steganography algorithm against statistical analysis," proceeding of the IEEE, vol. 19, pp. 3884-3888, August 2007.
- [5] R.J. Anderson, F.A.P. Petitcolas, "On the limits of steganography", IEEE Journal of selected Areas In Communications, vol 16, pp. 474-481, May 1998.
- [6] A. Cheddad, J. Condell, K. Curran, P. Kevitt, "Digital image steganography- survey and analysis of current methods," Signal Processing, vol. 90, pp. 727-752, 2010.
- [7] D. Bhattacharyya, A. Roy, P. Roy, T. Kim "Receiver compatible data hiding in color image," International Journal of Advanced Science and Technology, vol. 6, pp. 15-24, May 2009.
- [8] Adnan Gutub, Ayed Al-Qahtani, Abdulaziz Tabakh "Triple-A: secure RGB image steganography based on randomization" AICCSA, IEEE/ACS International Conference on Computer Systems and Applications, Rabat, Morocco, pp. 400-403, 2009.
- [9] A. Kaur, R. Dhir, G. Sikka, "A new image steganography based on first component alteration technique", International Journal of Computer Science and Information Security (IJCSIS), vol. 6, pp. 53-56, 2009.
- [10] M.T. Parvez, A. Gutub, "RGB intensity based variable-bits image steganography", APSCC 2008 -Proceedings of 3rd IEEE Asia-Pacific Services Computing Conference, Yilan, Taiwan, December 2008.
- [11] Nameer N. EL-Emam, "Hiding a large amount of data with high security using steganography algorithm," Journal of Computer Science, vol. 3, pp. 223-232, 2007.
- [12] Mohammed A.F. Al-Husainy, "Image steganography by mapping pixels to letters," Journal of Computer Science, vol. 5, pp. 33-38, 2009.
- [13] A. Ibraheem Abdul-Sada, "Hiding data using LSB-3," J.Basrah Researches (Sciences), vol. 33, pp. 81-88, December, 2007.

AUTHORS PROFILE



Sukhpreet Kaur received her B.Tech in Computer Science & Engineering from Punjab Technical University, Punjab, India in 2007. She is pursuing her M.Tech in Computer Engineering from Punjabi University, Patiala. There are more than 8 research papers in various national and international conferences in the credit of Ms Kaur. Her interest areas include fields of network security and steganography. Currently, she is working as lecturer in computer science in the Department of computer science and engineering, Baba Farid College of Engineering and Technology, Bathinda, Punjab State, India.



Sumeet Kaur received her B.Tech in Computer Engineering from Sant Longowal Institute of Engineering & Technology (Deemed University) Punjab in 1999 and her M.Tech from Punjabi University, Patiala in 2007. She has more than 10 research papers in different national and international conferences. Currently, she is working as lecturer in computer science in the Department of computer engineering, Yadavindra College of Engineering Punjabi University Guru Kashi Campus, Talwandi Sabo, Punjab State, India. Her interest areas include encryption, network security, image processing and steganography.

An approach to a pseudo real-time image processing engine for hyperspectral imaging

Sahar Sabbaghi Mahmoudi*
Smart Technology and Robotics Programme
Institute of Advanced Technology (ITMA),
Universiti Putra Malaysia, Serdang, Malaysia

Prof.Dr.Shattri Mansor
Remote Sensing and GIS Programme,
Department of Civil Engineering,
Universiti Putra Malaysia, Serdang, Malaysia

Abed Abedniya
MBA Programme,
Faculty of Management (FOM),
Multimedia University, Malaysia

Abstract

Hyperspectral imaging provides an alternative way of increasing the accuracy by adding another dimension: the wavelength. Recently, hyperspectral imaging is also finding its way into many more applications, ranging from medical imaging in endoscopy for cancer detection to quality control in the sorting of fruit and vegetables. But effective use of hyperspectral imaging requires an understanding of the nature and limitations of the data and of various strategies for processing and interpreting it. Also, the breakthrough of this technology is limited by its cost, speed and complicated image interpretation. We have therefore initiated work on designing real-time hyperspectral image processing to tackle these problems by using a combination of smart system design, and pseudo-real time image processing software. The main focus of this paper is the development of a camera-based hyperspectral imaging system for stationary remote sensing applications. The system consists of a high performance digital CCD camera, an intelligent processing unit, an imaging spectrograph, an optional focal plane scanner and a laptop computer equipped with a frame grabbing card. In addition, special software has been developed to synchronize between the frame grabber (video capture card), and the digital camera with different image processing techniques for both digital and hyperspectral data.

Keywords: *Remote sensing, image processing, Real-Time, frame grabber, hyperspectral, Hardware/Software Design.*

1. Introduction

Digital and Remote sensing image processing is nowadays a mature research area. Use of hyperspectral remote sensing in both research and operational applications has been steadily increasing in the last decade. Hyperspectral imaging systems can capture imagery from tens to hundreds of narrow bands in the visible to infrared spectral regions. These systems offer new opportunities for better differentiation and estimation of biophysical attributes and have the potential for identification of optimal bands and/or band combinations for a variety of remote sensing applications [1-3],[11]. Different remote sensing applications have proven to be potential sources of

reflectance data such as extraction of various vegetation spectral features. Satellite-based remote sensing provides a unique opportunity to obtain characteristics over large areas, whereas airborne remote sensing provides remotely sensed data over the medium scale, such as farms and small watersheds [4]. However, these studies largely depend on the availability of spectral images that are usually quite expensive and need to be acquired by professional image providers. Ground based hyperspectral imaging has been used as a cheap tool to acquire remotely sensed data from individual part of proposed area [4].

In this paper, we propose an approach to pseudo real-time image processing engine for hyperspectral imaging to increase mission flexibility for environmental planning, medical diagnostics, remote sensing, and natural resources applications. All processes in the implementation of hyperspectral imagery and remote sensing apply near real time image processing done at the spatial and numerical modeling laboratory (SNML) at the University of Putra Malaysia. The main focus of this research is the development of a camera-based hyperspectral imaging system for stationary remote sensing applications. Hyperspectral imaging provides an alternative way of increasing the accuracy by adding another dimension: the wavelength. Recently, hyperspectral imaging is also finding its way into many more applications, ranging from medical imaging in endoscopy for cancer detection to quality control in the sorting of fruit and vegetables.

The impetus in performing this research was given by existing snags and problems faced by workers in the field. So far, many of the image processing software available in the market do not process images in real time. The software has to download and read the images first and then prepare image-processing functionalities on them. In this paper, we attempt to show that it is possible to have pseudo-real image processing. This means that processing is done on the fly: as soon as the camera captures the image, the image processing algorithm comes into play immediately in all embedded applications.

* Responsible author

The Hyperspectral imaging system consists of four components:

- A sensing component: a hyperspectral sensor (high performance digital CCD camera now known as the ImSpector manufactured by SPECIM systems) for acquiring data or images.
- An optional focal plane scanner
- A video capture (frame grabber) card connected to the CPU from the camera helps in data capture.
- Acer Extensa Notebook 4630z, which is a 2.0 GHz Intel notebook, computer manufactured by Acer Inc., has been used as the CPU on the sensor part.

The remainder of this article is structured as follows. Section 2 presents essential characteristics and concepts in the scope of the work. Section 3 presents System Requirement. We will describe developing software in section 4. Description of proposed method, design and a relative technique are discussed in section 5. Section 6 shows the experimental result and discussions. Section 7 presents the conclusion of this paper.

2. Concepts and characteristics

In order to draw a clear picture of fundamental concepts and characteristics of hyperspectral imaging, it is important to recap some key concepts and definitions which are accepted by experts in this field.

Real-time image processing: operating systems serve application requests nearly real-time. In the other word manipulation of live images, typically within 50 to 100 milliseconds, so the human user perceives them as instantaneous.

Embedded systems: An embedded system is a computer system designed to perform one or a few dedicated functions often with real-time computing constraints. It is *embedded* as part of a complete device often including hardware and mechanical parts.

Engine: The image processing engine, or image processor, is an important component of a digital camera and plays a vital role in creating the digital image. The image processing engine comprises a combination of hardware processors and software algorithms. The image processor gathers the luminance and chrominance information from the individual pixels and uses it to compute/interpolate the correct color and brightness values for each pixel.

Pushbroom: In remote sensing, an imaging device consisting of a linear array of sensors (CCD camera) which is swept across the area of observation. Pushbroom

systems allow greater resolution of data to be assimilated than do line scanner systems.

3. Real-Time Hyperspectral Imaging System Requirement

3.1. The sensor

The hyperspectral sensor used in this study was a ground based user-friendly line sensor ImSpector (V10) See (Fig.1). The new ImSpector Fast10 is a high intensity imaging spectrograph, and makes spectral imaging possible at hundreds and even up to 1500 images per second. ImSpector Fast10 imaging spectrograph provides[5]:

- high light throughput
- superior image quality
- good spectral resolution of 15 nm
- full VNIR spectrum of 400 - 1000 nm over a narrow dimension, allowing short read out times
- maximum light intensity on the camera pixels, allowing short integration times
- high speed acquisition in many low cost industrial CCD and CMOS cameras

The ImSpector imaging spectrograph is a component that can be combined with a broad range of monochrome matrix cameras to form a spectral imaging device. Equipping the instrument with an objective lens coupled with a monochrome area camera, converts ImSpector to a spectral line imaging camera. Operation is based on the direct sight imaging spectrograph technology of the Spectral Imaging Ltd. (SPECIM), Oulu, Finland [6]. ImSpector captures a line image of a target and disperses light from each line image pixel to spectrum. Each spectral image then contains line pixels in the spatial axis and spectral pixels in the spectral axis (Fig. 2) [4]. It is possible to acquire full spectral information for each line image acquired from the target. Since ImSpector captures sequential images of the moving target (or the sensor itself moves), a 2D spectral image can be formed. This technology allows diverse opportunities to analyze the target accurately based on its spectral features.



Fig. 1– hyperspectral sensor (ImSpector V10).

3.1.1 Advantages of hyperspectral imaging system

Hyperspectral imaging is extremely advantageous in terms of its data, presenting the information in the spatial direction which is useful for extracting information with less loss of data. Some advantages of hyperspectral imaging over conventional techniques such as: NIRS (Near-infrared spectroscopy), RGB, and hyperspectral imaging are shown in Table 1 [7, 8].

Feature	RGB imaging	NIRS	MSI	HSI
Spatial information	✓		✓	✓
Spectral information		✓	Limited	✓
Multi-constituent	Limited	✓	Limited	✓
Sensitivity to minor Components			Limited	✓

Table.1 Advantages of hyperspectral imaging system

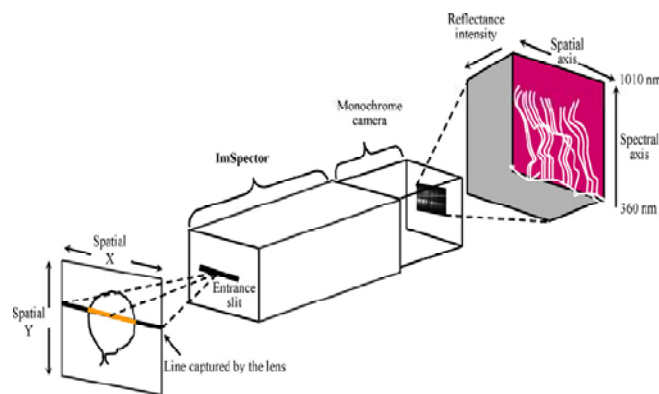


Fig. 2 – The operating principles of ImSpector.

3.2 An optional focal plane scanner

The focal plane scanner performs line scanning across an input imaging area within the focal plane of the front lens and the spectrograph disperses each line into a spectrum

and projects a two-dimensional image profile (line image) onto the CCD surface. This configuration allows image acquisition under stationary or laboratory settings [2], [10].

3.3 A video capture (frame grabber)

The FrameLink frame grabber is a TYPE II PC Card with both a Camera Link and Card bus interface. It provides the ability to capture digital video data from a 'base configuration' Camera Link interface and transfer that data to host memory via a Card bus (PCI) interface. The frame link is a professional state of the art PCMCIA card bus digital video capture card, allowing user to display, capture, store and preview mega pixel video image (up to 16 mega pixels) on the notebook computer [9]. The Imperx FrameLink video capture card is as shown in (Fig.3) below.



Fig.3 – The IMPERX FrameLink Fast CardBus video capture (frame grabber) card. This picture has been taken from the official website of Imperx Inc.

3.4 The computer system

The computer is an Intel Pentium III (800 MHz) processor based system with 250 GB hard drive. The operating system on the computer is Microsoft Windows XP. A PCI interface board provided with the imaging system is installed in a master PCI slot in the computer. The utility software is installed in the computer for complete camera control, image acquisition and applies image processing technique. The Acer Notebook computer is as shown in (Fig. 4) below.



Fig.4 – Different views of the Acer Extensa 4630z Notebook computer. This has been used as the CPU on our hyperspectral imaging system. These pictures have been obtained from Acer Inc.

3.5 Acquiring ground-based hyperspectral images

The ground-based hyperspectral line imaging systems is shown in (Fig. 5). The hyperspectral sensor ImSpector captures the scene. ImSpector captures a line image of the scene and disperses it into a spectrum. By moving the sensor up and down or left and right by means of a battery-powered movable tripod base, the whole scene is captured. The rate of image acquisition can be up to 30 fps, and data can be saved in an audio-video interleave (avi) file format. The raw spectral data obtained by the sensor and the image generated by applying a line pixel assembly algorithm to the raw data in an image [4]. Each frame represents the spectral data corresponding to a spatial line. The x-axis of each frame is the spatial axis and the y-axis is the spectral axis. Each frame is composed of 480 spectral lines, each representing spectral data at a particular wavelength. In order to facilitate comprehension of these spectral data, an image is generated by applying a line pixel assembly algorithm to every frame. Assembly of spectral lines with an equivalent wavelength from all frames makes one image, and thus the procedure can generate a total of 480 images, each displaying the scene captured with a different wavelength [6].

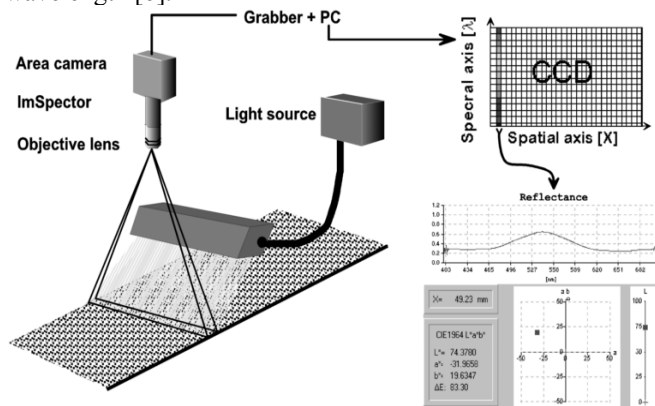


Fig. 5 – Hyperspectral image acquisition system.

4. Developing the software

Once the hyperspectral images are generated, they would appear as a stack of continuous images. Manipulation of hyperspectral images and extraction of useful spectral information from these multidimensional data requires the development of intelligent software. For this purpose, software with many high level computing and visualization functions embedded in a number of useful toolboxes. (Fig.6). illustrates the main menu and its user interfaces for image processing and data extraction.

Furthermore, the functions embedded in the popup menu endow the interface with many other image processing and parameter extraction capabilities.

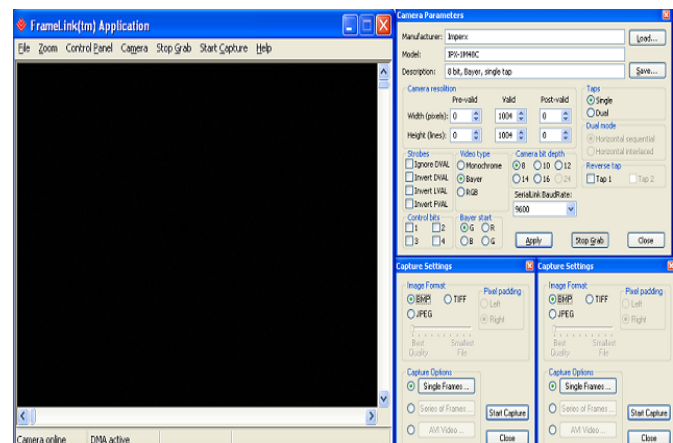


Fig. 6 – Software interface

The Image Processing Toolbox provides a comprehensive set of standard algorithms and graphical tools for image processing, analysis, visualization, and algorithm development. You can restore noisy or degraded images, enhance images for improved intelligibility, extract features, analyze shapes and textures, and register two images. Most toolbox functions are written in the C++ language. A schematic diagram of the interface design and its utilities is shown in (Fig. 6, 9).

4.1 Some key features of image acquisition toolbox

- Image enhancement, including linear and nonlinear filtering, filter design, and automatic contrast enhancement;
- Binarization filters (threshold, threshold with carry, ordered dithering, Bayer dithering, Floyd-Steinberg, Burkes);
- Color image processing, including color space conversions and channel replacing, channel filtering;
- Spatial transformations and image registration, including a graphical tool for control-point selection;
- Fourier transformation (low pass and high pass filters);
- Mathematical morphology filters (erosion, dilatation, opening, closing, hit & miss, thinning, thickening);
- Edge detectors (homogeneity, difference, sobel, canny);
- Median filter, Adaptive smoothing, Conservative smoothing;

5. Methods

In this technique only hyperspectral data for areas of interest captured. The principal component analysis method is used to hyperspectral imaging systems shown in (Fig.8.)

Based on the raw data, the software developed is used to generate images for a three-dimensional, including two spatial axes and one spectral axis that can be produced in one of four ways: Fourier transform imaging, point-to-point spectral scan in a spatial grid pattern, line by line spatial scan i.e. the pushbroom technique, and wavelength tuning with filters. The line by line spatial scan and wavelength methods are more suitable. In the pushbroom method spectral data acquire across the full spectral range of single spatial lines consecutively to reconstruct the hyper spectral tube.

The CCD camera provides 1280(h) x 1024(v) pixel resolution and true 12-bit dynamic range. The imaging spectrograph is attached to the camera via an adapter to disperse radiation into a range of spectral bands. The effective spectral range resulting from this integration is from 400 nm to 1000 nm. Diffuse illumination of the sample is made possible by a florescent-halogen or LED source [6]. The light reflected from the target enters the objective lens and then spread into its component wavelengths as shown in (Fig7).

The optional focal plane scanner can be attached to the front of the spectrograph via another adapter for stationary image acquisition. The camera and the frame grabbing card are connected via a double coaxial cable, and the utility software allows for complete camera control and image acquisition. The imaging system captures one line image for all the bands at a time and the focal plane scanner serves as a mobile platform to carry out pushbroom scanning in the along direction.

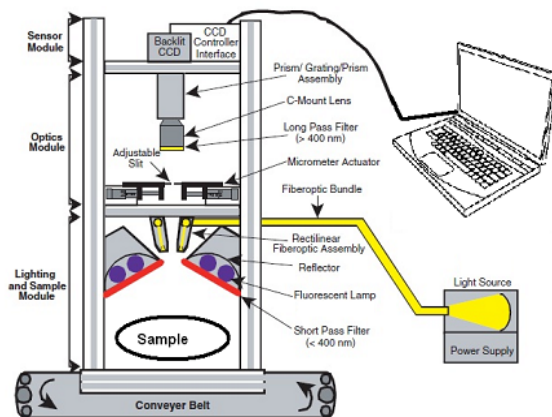


Fig 7– A Scheme diagram the of current Hyperspectral imaging system

Preliminary image acquisition testing trials indicate that this CCD camera-based hyperspectral imaging system has potential for agricultural and natural resources applications. (Fig.8) shows the architecture of the ground-based hyperspectral line imaging systems based on architecture at SMNL laboratory in UPM.

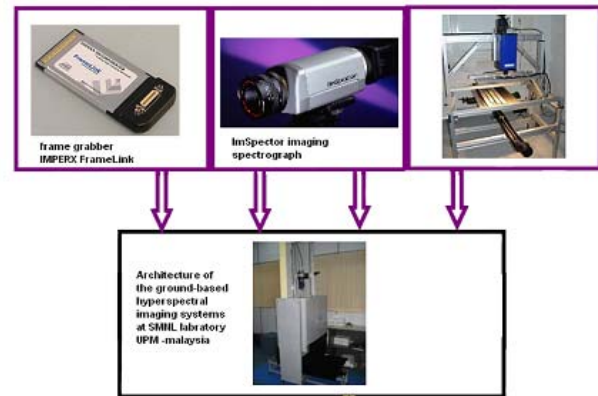


Fig. 8 – architecture of the ground-based hyperspectral imaging systems

6. Experimental result and discussion

In order to test the sensor design concept and to integrate software design, we simulate a realistic scene. The Digital Imaging and hyperspectral software, developed at Institute of Advance Technology (ITMA) in Malaysia. By scene simulation and sensor modeling, we hope to reduce the cost and development time in new sensor designs, together with the support of the algorithm and the techniques method. The image processing algorithms are designed only to demonstrate the idea of effectively capturing hyperspectral data. Needless to say, more sophisticated algorithms need to be developed for more challenging tasks.

After software executed, the main window will appear. The main window provides the primary area for viewing real-time images received from the camera. When image viewing is active, pull-down menu with two options reveals: 'Player' and 'Exit'. Player button will toggle between 'Start Grab' and 'Stop Grab' every time the user clicks on it. By clicking on 'Start Grab' enables the engine and causes the main window to display live images received from the camera. Clicking on 'Stop Grab' disables the engine and causes the display to freeze. When recording images to disk, Image Format option selects the format, 'BMP', 'JPEG' or 'TIFF' that the image will be saved in. Selecting 'JPEG' activates a compression slider. 'Best Quality' provides the least compression while 'Smallest File' provides the most compression.

In order to evaluate the system and simulate a realistic scene we pluck the leaf from the tree in the fields near our campus and around University of Putra Malaysia. In hyperspectral imaging system design, different portion of bandwidth can be selected and determined by analyzing model spectral profile combined to a single image profile and a binary decision was made using a threshold found by experience. Thus, object can be demonstrated in real time. In (Fig. 9) we show a single image snapshot captured, and the result was combined to produce a co-registered composite image. After capture the scene raw data save as 'JPEG' format then we apply some image processing technique in order to assess our software.

Here we select histogram filter to determine the overall intensity of the image that is suitable for our inspection task. Based on the histogram data, you can adjust your image acquisition conditions to acquire higher quality images see (Fig.9).

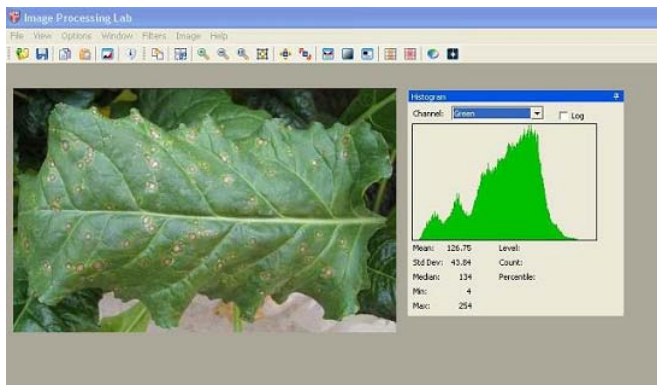


Fig.9 – captures the scene and apply Histogram Filter

Another filter that applied for evaluate our system work is thresholding (Fig. 10). Objective of thresholding filter is converting the image into binary objects. Thresholding is the simplest method of image segmentation. From a grayscale image, so we could apply the basic morphology processes, Image analysis capability can also be expanded to include other types of analytical techniques for a particular image analysis purpose.

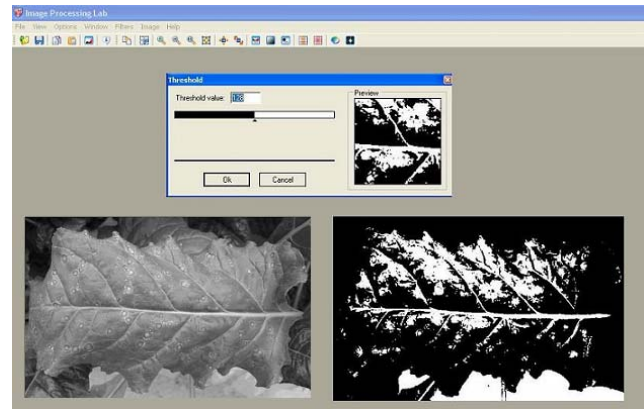


Fig.10 – Apply Thresholding Filter

The hyperspectral line sensor captures the raw reflectance data in the approach that illustrated above. As this is a ground-based system, the cost is much lower than for airborne- or satellite-based remotely sensed data. The nominal spectral resolution of 1.5–2nm within the wavelength range of 400–1000nm is sufficient for most application studies. The software developed serves a pivotal role in dealing with the spectral data that are captured. It can generate images from the raw spectral data in an audio–video interleave format or image format. Useful image analysis algorithms are included, such as; Thresholding and other functions determine whether an image meets certain criteria for inclusion in an analysis.

7. Conclusions

This paper reviews the recent developments in ground-based hyperspectral imaging system for acquisition of reflectance data that is useful for many real-life applications such as; environmental planning and natural resources applications. The hyperspectral imaging technique described in this article provides a new opportunity for determining the optical properties and quality of product such as food and agricultural products.

Compared to other techniques, the hyperspectral imaging technique is simpler, faster and easier to use, and more importantly it is capable of determining optical properties over a broad spectral range simultaneously. The technique also is useful for measuring the optical properties of turbid food and agricultural products. Moreover the hyperspectral imaging technique is potentially useful in assessing, sorting, and grading fruit quality.

References

- [1] D.Tuia, and G.Camps-Valls, "Recent advances in remote sensing image processing", IEEE International Conference on Image Processing (ICIP), 2009 , 3705 – 3708.
- [2] H. James Everitt, and R. Michael Davis , and Chenghai Yang, "A CCD Camera-based Hyperspectral Imaging System for Stationary and Airborne Applications" , Geocarto International, Vol. 18, No. 2, June 2003.
- [3] H. James Everitt, Chenghai Yang, Joe M. Bradford and Dale Murden, Airborne Hyperspectral Imagery and Yield Monitor Data for Mapping Cotton Yield Variability, Volume 5, Number 5, 445-461,2004.
- [4] Xujun Yea, Kenshi Sakaib, Hiroshi Okamoto, Leroy O. Garcianod," A ground-based hyperspectral imaging system for characterizing vegetation spectral features",Elsevier Science Publishers B. V, Volume 63, NO 1, 13-21 , August 2008.
- [5] Official website of Spectral Imaging Ltd, Finland.
SPECIM: <http://www.specim.fi/>
- [6] Users Manual for Inspector spectrograph Ver.2.0 from SPECIM website.
- [7] A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., and Frias, J. M. 2007. "Hyperspectral Imaging - an Emerging Process Analytical Tool for Food Quality and Safety Control", Volume 18, Issue 12, 2007, 18(12); 590-598.
- [8] Osama M. Ben Saaed, Abdul Rashid Mohamed Shariff, Helmi Zulhaidi Mohd Shafri, Ahmad Rodzi Mahmud, Meftah Salem M Alfathi, "Hyperspectral Technique System for Fruit Quality Determin ", Map Asia 2010 and ISG 2010 Conference, July 2010.
- [9] Official website of Imperx Inc, USA .
FrameLink: <http://www.imperx.com/>
- [10] C. Mao., "Hyperspectral imaging systems with digital CCD cameras for both airborne and laboratory application", 17th Biennial Workshop on Videography and Color Photography in Resource Assessment, American Society for Photogrammetry and Remote sensing, Bethesda, MD. pp. 31-40,1999.
- [11] C. Mao., " Hyperspectral focal plane scanning-an innovative approach to airborne and laboratory pushbroom hyperspectral imaging. Proc. 2nd International Conference on Geospatial Information in Agriculture and Forestry", ERIM International, Inc.,Ann Arbor, MI. Vol. 1, pp. 424-428, 2000.

AUTHORS PROFILE



Sahar Sabbaghi Mahmouei is currently doing her Master degree in Institute of Advanced Technology and Research (ITMA), Universiti Putra Malaysia, UPM. Sahar has received her B.Sc in software computer engineering field in 2006 from Iran Azad University. Her research interest includes image processing, machine vision, artificial Intelligence and e-commerce.

Improved Computer Networks Resilience Using Social Behavior

Yehia H. Khalil^{1,2}, Walaa M. Sheta¹

¹Dept. of Virtual Reality & Computer Graphics
Informatics Research Institute, MuCSAT
New Borg El-Arab, Egypt

Adel S. Elmaghraby

²Dept. Computer Science and Computer Engineering
University of Louisville
Louisville, KY, USA

Abstract— Current information systems face many challenges in terms of malicious activities, hacking threats and natural disasters. A key challenge is to design a resilient communication networks that can provide high performance level with minimum disconnected points and delay. This paper presents a novel approach to discover the most critical network's nodes based on social network analysis (SNA) which have been used for social studies and recently have been widely used in many domains. The main focus of social network analysis is to study the "relations" between network nodes. In principle, critical network's nodes will be identified based on their magnitude for the network in terms of centrality: Degree, Betweenness and Closeness. The results show that using social network analysis enhances computer network resilience by identifying the critical elements of communication network. (Abstract)

Keywords- Network Resilience; Social Network Analysis; Redundancy; Critical Cyber Infrastructure. (key words)

I. INTRODUCTION

Computer networks are a combination of several resources: software, hardware and others. A resilient computer network refers to the ability of the network to operate and provide services with minimum delay under sever operational conditions and the ability to recover of the failure within acceptable time range. A diversity of failures which can cause local or wide disconnection, failures used to be caused by downtime of devices or miner power outage, yet other categories such as natural disasters, and malicious activates both for software or hardware elements had been added to the list [1].

Resilience is a very basic and significant prerequisite for any information system, the term Resilience has been defined in several domains. The main characteristics of any resilient system are: continuity of service delivery under unexpected operational environment and speed recovery from any failure [2]. Subsequently, a resilient computer network is the network which provides high data transfer rate with minimum latency. Building a resilient computer network involve several considerations: identifying critical elements, provide alternatives, develop recovery policies and mechanisms.

The main focus of this work is to investigate the use of social network analysis for identifying computer networks critical elements. So network managers, planner and

designers can develop better recovery plans and selected elements to be redundant.

Social network is a social construction made of nodes such as individuals, organizations, etc. which are called "nodes/stars" which are attached to each other by one or many connections/relations such as financial exchange, similar interest, common experiences, etc. [3] the following Fig.1 illustrates a simple social network.

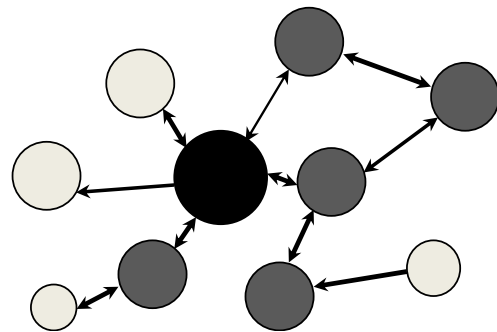


Figure 1. Simple Social Network Example

As shown, each entity can be connected to any number of entities; also the relation between the entities can be one direction or bi-direction based on the relations type. [4]

The conventional data representations differ on social networks in several aspects such as: the focus of social network is the relation between elements and the ability to build layers based on the amount of details targeted to be study, for example the relation between the government organizations can be represented by one network and another layer can represent the relation between the departments within the same organization. In addition, figure 1 illustrates the relations between the entities are very clear in terms of level of the relations as visualized by the link width also which entity are connected to all or some of the nodes which reflects the entity magnitude which highlight the difference between the difference between the social network representation and other data representations.

Social network analysis has been used for several years as analysis tool of social relation between humans in the sociology domain as humans tend to group based their interest, experience and etc.

II. BACKGROUND

This section demonstrates the computer networks resilience and social network analysis concepts.

A. Computer Networks Resilience

As a matter of fact, everything in our life relies on communications and computer networks represent a large portion of it. Currently our computer networks are not resilient or secure enough to face the strategic vulnerabilities of the nation [5]. Generally, the computers, servers, etc which are connected to the network are attacked for several reasons. However, using certain techniques such as using virtual privet network (VPN), encryption, firewalls, and other techniques can eliminate many of those attacks. Yet hackers, natural disasters can target other network elements such are routers, in many cases hackers amid to shutdown routers to develop discontinuity holes or developing malicious activities affect network performance and security [6].

A major step in building resilient network is identifying the critical elements which would need extra attention and secure them using the appropriate techniques. Network managers, designer and planners tend to use redundancy to avoid network failures. The following figure (Fig. 2) illustrates how redundancy can enhance network resilience.

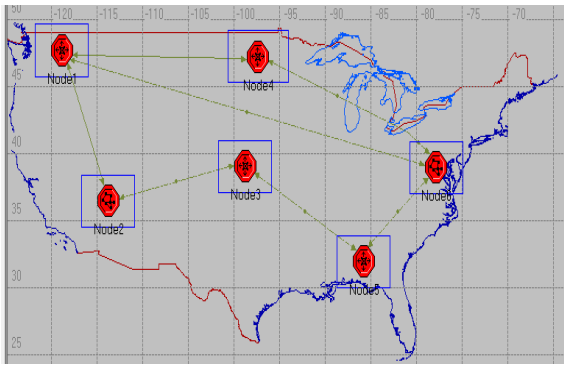


Figure 2. Network Example

As Fig. 2 shows, there are six nodes connected through seven links. At any time one of the links or the nodes can be down, the best way to ensure resilience is to redundant all the server and create a full connect network (Mesh). Unfortunately, that would be a very expensive solution not many can afford. The second approach is to determine the most important critical elements which need to be redundant and develop policies and algorithms for speedy activation for the backup devices.

B. Social Network Analysis

Social network analysis is an emerging set of techniques and schemes for data analysis, many researchers and scientists introduced several definitions based on their domain of interest. For example Hannemann proposed: "A social network is a set of actors that may have relationships with one another. Networks can have few or many actors (nodes), and one or more kinds of relations (edges) between pairs of actors." [7]

The major deviation for social network analysis over the other traditional approaches is its focus is to analyze information based on the relation between data entities. Social network can be represented as matrices or graphs; the plus of using graph is the ability to represent different types of relations between the nodes. One of the important concepts of social networks analysis is the hierarchical analysis, as the analysis can be proceed on different levels: Node, Dyadic, Triadic, Subset, and Network level [8]. However, the focus of the majority of research work is narrowed to the node and network level.

At network level, network density can be obtained by dividing the number of relations by the number all possible relations, the result varies between 0 and 1 and the higher ratio the denser network [8]. Another level would be the node level where it more concern about how important is the node? How popular is the node? Is it a central node?

Within the context of social networks the term power/centrality refers to the impact of this node on others nodes, and what would be the consequence in case of removing this node. Social network analysis offers three measurements for centrality: Degree centrality, Closeness centrality and Betweenness centrality [9] [10].

Degree Centrality: the degree centrality of a node A (DC_a) is number of connections/relations the node has. The node/actor with higher number of relations or ties maintains a higher traffic (in/out).

$$DC(N_i) = \sum_{j=1}^n a_{ij}$$

Where:

$DC(N_i)$: Degree Centrality of node N_i ,
 A : an adjacent matrix of relations network,
 n : number of nodes.

Centrality closeness: indicates how a node N_i close to the other nodes, depending on the application closeness would have different ways to be calculated. In computer networks scenario, our target will be physical distance.

$$CC(N_i) = 1 / \sum_{j=1}^n d(N_i, N_j)$$

Where:

$CC(N_i)$: Closeness of N_i ,
 $d(N_i, N_j)$: absolute distance between node N_i and node N_j ,
 n : number of nodes.

Centrality Betweenness: it measure characterizes of nodes as having a powerful positional i.e. a node is frequently shown in communication paths between any other nodes.

$$CB(N_i) = \sum_{j,k} \frac{P_{j,k}(N_i)}{P_{j,k}}$$

Where:

$CB(N_i)$: Betweenness of N_i ,
 $P_{j,k}(N_i)$: shortest path between N_j, N_k and has N_i on it
 $P_{j,k}$: shortest path between N_j, N_k

The following section will illustrate the network resilience problem and the proposed approach to enhance it.

III. PROBLEM STATEMENT

Computer networks are the core on any information system, the ability of any network to maintain an acceptable service level throughout malicious activities is called resilience [5]. Anchored in resilience definition, building a resilient computer network consolidates two main aspects:

- I. Device's redundancy: installing backup devices, such as power supplies, routers, switches, etc. that kicks in when the primary fails.
- II. Develop recovery methodologies and Policies: how to use the backup systems to ensure minimum quality of services (QoS) variation in case of emergency.

Vendors will tell that we need to go with full redundancy which is great but that will requires large investments and also is complex for monitoring or management purposes. Therefore selecting the critical elements to be redundant is a vital process, calculating the probability of system failure is one of the well known approaches for redundancy as the more duplication the less failure probability [11].

By enlarge this approach have several drawbacks such as: assuming failure independency and non-realistic estimation of different probability weights. The uses of social network analysis provide more realistic information about nodes importance and consider the correlation between devices failure.

Several routing and networking parameters can be affected when one of the routers fail down such as network latency, routing tables size, and packet drop rate. In this study we will focus on network latency as it can reflect the overall network performance. The failure of a critical router or node should cause a huge change on network latency, so with no backup devices installed scenario the methodology is to evaluate network latency change to validate the social network analysis approach.

IV. EXPERIMENTS

The main purpose of those experiments it is to validate the ability of social network analysis methods at identifying critical routers within a network.

Experiments configuration: for illustration purpose, the simulation scenarios were based on a modified version of the University of Louisville computer routers infrastructure as shown in the Fig. 3 [12]. The physical topology was imported to the OPNET simulation tool, also network traffic were collected between network routers and exported to the simulation tool.

Testing scenarios: For testing purpose, malicious actives were simulated either by injecting the system with overloading traffic or implementing a node failure. A traffic broadcasting node was hocked up to the network to implement both scenarios, the traffic generation process follows the Exponential distribution with $\lambda = 0.025$ and 1 as shown on Fig 4.

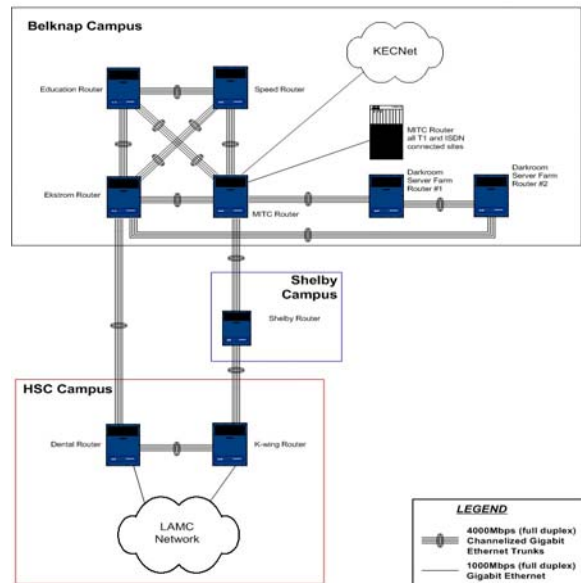


Figure 3. University of Louisville Gigabyte Backbone
(Source: Miller Information Technology Center, U. of Louisville)

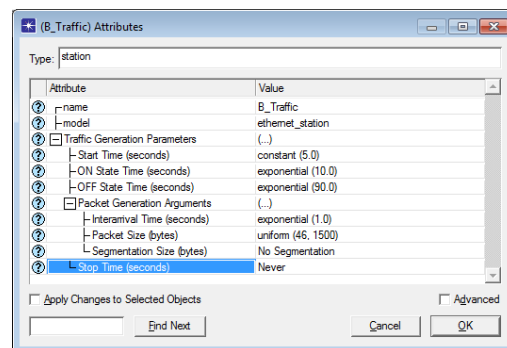


Figure 4. Traffic Generation Node Configurations

For validation purpose, simulation will run with two routers fail/recovery scenario and network latency information will be collected. As shown in the following figures: Fig. 5 shows the modified network and Fig. 6 shows network latency has two cases of variation (A, B) although that the failed/recover routers have the same capacity/configuration/manufactures, it was shown that each one has affected the network latency differentially.

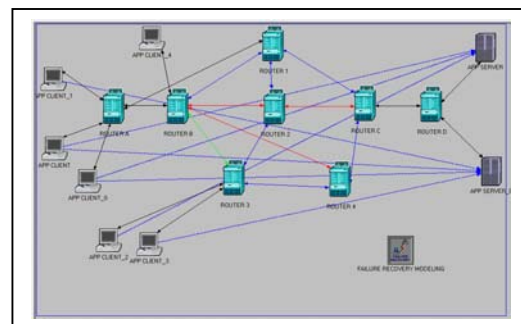


Figure 5. Network Topology

V. RESULTS AND DISCUSSIONS

Network ties visualization represent easier way to understand network behavior as shown on the following figures: Fig.9 and Fig. 10.

For example, by visual inspection, it is clear that some nodes process higher traffic than others also that routers A & D can be identified as sources/distention points. The first step in Social network analysis is to calculate the network density, this information can be used to determine the possibility for adding more paths/connection between nodes with constrain to the hardware limits. The calculated Network Density= 0.4642857 and Weighted Network Density = 87.41071, which indicated that adding more nodes or links to accommodate more traffic and services.

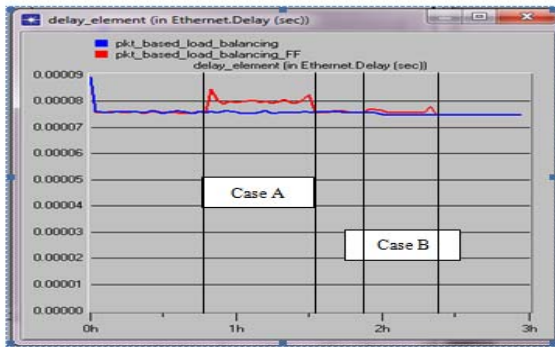


Figure 6. Network Latency

For social network analysis, The Applied Graph and Network Analysis (AGNA 2.1); an application in use in for communication networks analysis [13]. The following graph represents the sociomatrix; a matrix of size (8×8) represents the ties between network elements. For comparison and validation purposes, we build two sociomatrix as shown in the following figures: Fig. 7 and Fig. 8

- 1-Uniform sociomatrix: all the links have the same weight and symmetric matrix.
- 2-Weighted sociomatrix: each link got its wight based on the throughput rate bits/sec in average created nonsymmetrical matrix.

	A	1	2	3	4	5	6	D
A	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
1	1.0	0.0	1.0	1.0	1.0	0.0	0.0	0.0
2	1.0	1.0	0.0	1.0	0.0	1.0	1.0	0.0
3	0.0	1.0	1.0	0.0	1.0	1.0	0.0	0.0
4	0.0	1.0	0.0	1.0	0.0	0.0	1.0	1.0
5	0.0	0.0	1.0	1.0	0.0	0.0	1.0	0.0
6	0.0	0.0	1.0	0.0	1.0	1.0	0.0	0.0
D	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0

Figure 7. Uniform Network Sociomatrix

	A	1	2	3	4	5	6	D
A	0.0	125.0	126.0	0.0	0.0	0.0	0.0	0.0
1	125.0	0.0	125.0	631.0	125.0	0.0	0.0	0.0
2	125.0	125.0	0.0	125.0	0.0	125.0	250.0	0.0
3	0.0	125.0	123.0	0.0	634.0	125.0	0.0	0.0
4	0.0	125.0	0.0	125.0	0.0	0.0	125.0	0.0
5	0.0	0.0	125.0	125.0	0.0	0.0	125.0	125.0
6	0.0	0.0	125.0	0.0	125.0	125.0	0.0	0.0
D	0.0	0.0	0.0	0.0	631.0	0.0	0.0	0.0

Figure 8. Weighted Network Sociomatrix

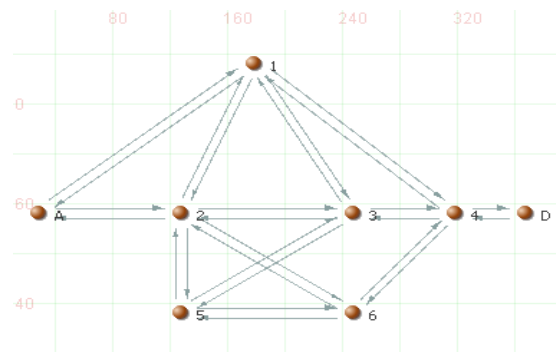


Figure 9. Uniform Network Visualization

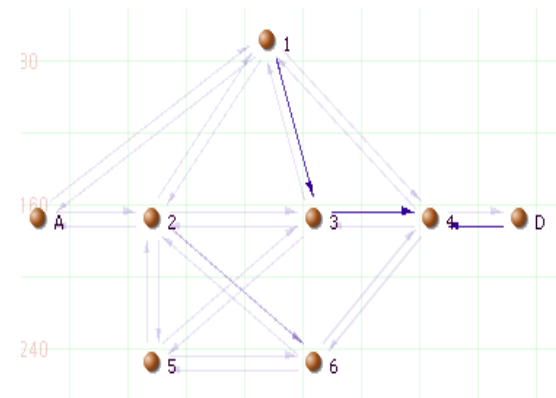


Figure 10. Weighted Network Visualization

The following step is to evaluate the Centrality based on the physical layout and concoctions; the results show no difference between the uniform networks and the weighted network which match the logic of those metrics.

The ANGA tool calculates the Centrality/Degree entitled Nodal Degree. The following table represents the nodal degree for each node and also compares it to other nodes.

TABLE I. DISTRIBUTION OF NODAL DEGREE

Node	Degree	Relative Degree*	Relative Degree**
A	2.0	0.285714285	0.25
1	4.0	0.571428571	0.5
2	5.0	0.714285714	0.625
3	4.0	0.571428571	0.5
4	4.0	0.571428571	0.5
5	3.0	0.428571428	0.375
6	3.0	0.428571428	0.375
D	1.0	0.1428571428	0.125

* Relative to number of all other nodes (self excluded) (Table footnote)

** Relative to number of all nodes (self included) (Table footnote)

As the table (Table 1) shows, routers (in order): 2, 3, 4 and 1 have higher level of centrality/degree as those nodes have higher number of relations which provide more flexibility. The next step is to evaluate Centrality/ Betweenness on a node level; ANGA 2.1 generates the following table.

TABLE II. DISTRIBUTION OF BETWEENNESS CENTRALITY

Node	Betweenness
A	0.0
1	6.3333335
2	7.6666665
3	4.3333335
4	13.666667
5	0.6666667
6	3.3333333
D	0.0

As shown routers A and D have the lowest Betweenness level, the network was designed as router A and D are source and destination points which confirm the obtained results. In addition, router 4 has the highest level and that confirmed as it the only router connected to destination point. Routers: 4,3,2,1 have higher level of betweenness.

TABLE III. DISTRIBUTION OF CLOSENESS CENTRALITY

Node	Closeness
Router A	0.07692308
Router 1	0.1
Router 2	0.1
Router 3	0.1
Router 4	0.1
Router 5	0.083333336
Router 6	0.09090909
Router D	0.0625

The last measurement of Centrality of is Centrality/Closeness; this index is the inverse of the sum of the geodesic distances from that node to all the other nodes as follow. It provides vital information for network planning and design concern.

By excluding router A and router D as they are the source and destination, we can see that routers 4, 2, 1, and 3 are very close to other nodes.

SNA concluded that router B, router C and router 1 are the most critical element in this network, for validation purpose the next set of experiments will examine how those nodes failure will impact the network performance in terms of latency and throughput rate to the destination nodes.

In this scenarios several routers (designated routers by SAN and others ones) will be failed and recovered, this will be done in isolated scenarios with the same starting failure time and for the same time period. The following figure demonstrates the Fail/recover procedure for each router.

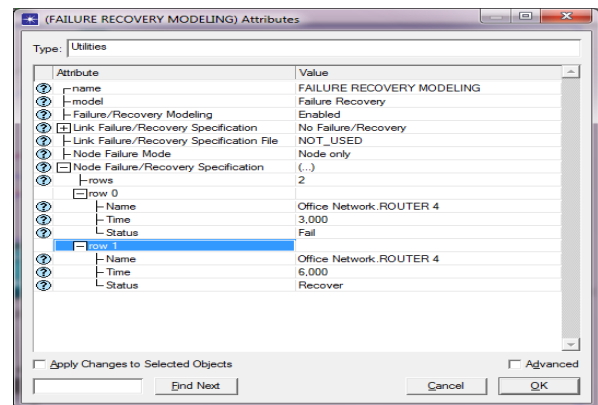


Figure 11. Fail/Recover router setting

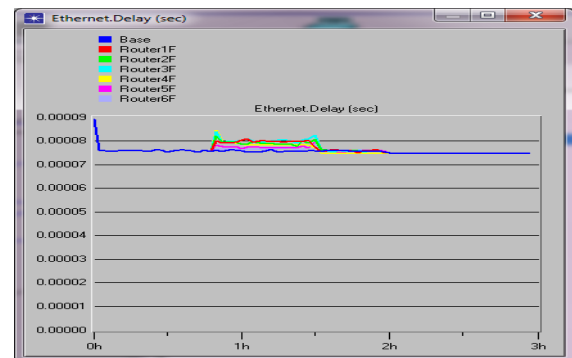


Figure 12. Global Ethernet Delay for the Routers Failure Time

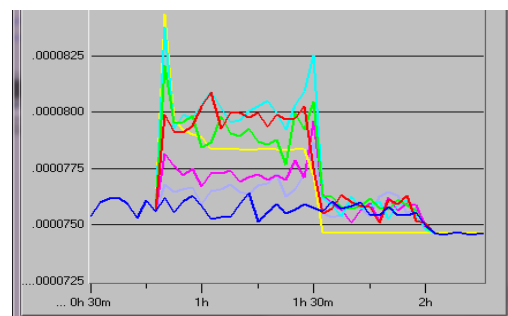


Figure 13. Zoomed Section for the Routers Failure Time

Fig. 12 and Fig.13 represent the global delay for the network, as shown each router failure impacted the latency differentially. SNA concluded that routers 4, 3, 2 and 1 are critical/central elements which are confirmed as they caused higher level of latency.

The Centrality measurements: Degree, betweenness and Closeness identified the critical elements of the network, for budget planning it is important to order the router based on their criticalness or importance over the network. To study

the nodes status in regard edges weight, the ANGA 2.1 sociometric status.

The following tables table illustrates sociometric status for each node within the uniform and weighted network.

TABLE IV. SOCIOMETRIC STATUS FOR EACH NODE (UNIFORM/WEIGHTED NETWORK)

	Uniform Network	weighted network
Node	Status	Status
A	0.5714286	71.57143
1	1.1428572	268.7143
2	1.4285715	196.28572
3	1.1428572	423.2857
4	1.1428572	287.85715
5	0.85714287	189.28572
6	0.85714287	189.28572
D	0.2857143	125.0

The obtained results for the sociometric show that for both cases (uniform/weighted) the routers 4, 3, 2 and 1 have higher weight than the other elements, however within the uniform network it hard to identify the order of their importance. While for the weighted network case, the critical router can be order to select the most important one.

In this work a small network was used for demonstration purpose, the Social Network Analysis designated the critical and important routers based on their Centrality evaluation. For validation purpose, network performance parameter: network latency was evaluated. The results showed that the Social Network Analysis successfully identified the critical rescuers of the investigated network.

VI. CONCLUSION

This research work presents a novel approach for identifying critical elements of computer networks consequently the network designers, planners and administrators can come to a decision regarding which elements should have recovery devices as a step toward enhancing the network resilience level. Social Network Analysis identifies the critical elements based on Centrality measurements for uniform and weighted networks; Sociomatrix provides flexible representation to accommodate various networks connection/edges strength and direction. The illustrated results showed that SNA successfully designated the critical routers, in addition SNA can provide vital information for network design process such as: shortest path, etc.

REFERENCES

- [1] Alan T. Murray and Tony H. Grubestic: "Overview of reliability and vulnerability in critical infrastructure", Springer Berlin Heidelberg, 2007. (references)
- [2] Kishor S. Trivedi, Dong Seong Kim, Rahul Ghosh: "Resilience in computer systems and networks", International Conference on Computer Aided Design Proceedings, California, 2009, pages: 74-77.
- [3] Linton Freeman: "The development of social network analysis", Vancouver, Empirical Press, 2006.
- [4] Hanneman and M. Riddle: "Introduction to social network methods", online: <http://www.faculty.ucr.edu/~hanneman/nettext/>, 2005.

- [5] Cyberspace Policy Review: "Assuring a trusted and resilient information and communications Infrastructure", White House Policy Review, 2009.
- [6] Desmedt, Y.: "Unconditionally private and reliable communication in an untrusted network", IEEE Information Theory Workshop on Theory and Practice in Information-Theoretic Security, 2005, pages: 38-41.
- [7] Hannemann R.A.: "Introduction to social network methods", <http://faculty.ucr.edu/%7Ehanneman/SOC157/TEXT/TextIndex.html>, 2001.
- [8] Wasserman, S. and K. Faust: "Social network analysis", Cambridge University Press, 1994.
- [9] Freeman, L.C., Borgatti, S.P. & White, D.R.: "Centrality in valued graphs: A measure of betweenness based on network flow", Social Networks, Vol.13, Issue 2, 1991, pages:141-154.
- [10] John Scott: "Social network analysis", London, Sage Publications, 1987.
- [11] DC Connors: "The variability of system failure probability", Reliability Engineering journal by Elsevier Ltd., Volume 9, Issue 2, 1984, pages: 117-125.
- [12] <https://www.it-comm.louisville.edu/router/>
- [13] Benta M: "Studying communication networks with AGNA 2.1", Cognition Brain Behav. 9, 2005, pages:567-574.
- [14] Rong Yang and Adel S. Elmaghraby: "A graph model for generating simulated web domains", 16th International Conference on Software Engineering and Data Engineering (SEDE-2007), USA, pages: 259-264.
- [15] Abdul Jabbar Mohammad, David Hutchison, and James P.G. Sterbenz, "Poster: Towards quantifying metrics for resilient and survivable Networks", 14th IEEE International Conference on Network Protocols (ICNP 2006), Santa Barbara, California, USA, November 2006.
- [16] E. Bursztejn and J. Goubault-Larrecq: "A logical framework for evaluating network resilience against faults and attacks.", 12th annual Asian Computing Science Conference (ASIAN), Springer-Verlag, Dec. 2007, pages 212-22.
- [17] Chigo Okonkwo; Martin, R.; Ferrera, M.P.; Guild, K.; O'Mahony, M.; Everett, J: "Demonstration of an application-aware resilience mechanism for dynamic heterogeneous networks.", International Conference on Transparent Optical Networks, 2007, Pages 20-23.
- [18] Peter J. Carrington, John Scott, Stanley Wasserman: "Models and methods in social network analysis", Cambridge University Press, 2005.
- [19] Maryann M Durland; Kimberly A Fredericks: "Social network analysis in program evaluation", Jossey-Bass, 2006.
- [20] Srivastava, J.: "Data mining for social network analysis", IEEE International Conference on Intelligence and Security Informatics, 2008, pages: xxxiii-xxxiv.
- [21] Jamali, M., Abolhassani, H.: "Different aspects of social network analysis", International Conference on Web Intelligence, 2006, pages: 66-72.
- [22] Uliuru, M.: "Design for resilience of networked critical infrastructures", Digital EcoSystems and Technologies Conference, 2007, pages: 540 - 545.
- [23] Haixin Duan; Jianping Wu: "Security management for large computer networks", Fifth Asia-Pacific Conference on Communications, 1999, pages: 1208-1213.

AUTHORS PROFILE

Yehia H. Khalil



He received B.Sc. with major of computer science and statistics from the Faculty of Science, University of Alexandria in 1994 and received M.Sc. with major in computer science and operations research from the Arab Academy for Science and Technology, Alexandria, in 2001. Yehia Khalil is a graduate student at the Computer Engineering and Computer science department, University of Louisville. He worked as a Researcher Assistance at Informatics Research Institute at Mubarak city for Scientific Research (MUCSAT) since 2001. His research interest includes data centre resilience, cyber-infrastructure security, computer networks performance and health outcomes research.

Walaa M. Sheta



He received his M.Sc. and PhD in Information Technology from the Institute of Graduate Studies and Research University of Alexandria, in 1992 and 2000, respectively. He received B.Sc. from the Faculty of Science; University of Alexandria in 1989. He is an associate professor of Computer graphics in Informatics Research Institute at Mubarak city for Scientific Research (MUCSAT) since 2006. During 2001-2006 he has worked as Assistant professor at MUCSAT. He holds a visiting researcher position at University of Louisville in US and University of Salford in UK. His research interest includes virtual reality, real-time computer graphics, Human computer Interaction and Scientific Visualization. He received M.Sc. and PhD in Information Technology from Institute of Graduate Studies and Research, University of Alexandria, in 1992 and 2000, respectively. He received B.Sc. from Faculty of Science, University of Alexandria in 1989.

Adel S. Elmaghraby



Adel S. Elmaghraby is Professor and Chair of the Computer Engineering and Computer Science Department at the University of Louisville. He has also held appointments at the Software Engineering Institute - Carnegie-Mellon University, and the University of Wisconsin-Madison. His research contributions and consulting spans the areas of Intelligent Multimedia Systems, Networks, PDCS, Visualization, and Simulation. He is a well-published author (over 200 publications), a public speaker, member of editorial boards, and technical reviewer. He has been recognized for his achievements by several professional organizations including a Golden Core Membership Award by the IEEE Computer Society. He is a senior member of the IEEE, a member of ACM and ISCA. He served a term as an elected ISCA Board member and currently is a Senior Member and an Associate editor for ISCA Journal. He is also senior member of the IEEE and a member of the IEEE-CS Technical Activities Board.

Mobile Embedded Real Time System (RTTCS) for Monitoring and Controlling in Telemedicine

By

Dr. Dhuha Basheer Abdullah
Asst. Prof./computer sciences Dept.
College of Computers and Mathematics/
Mosul University
Mosul/Iraq

Dr. Muddather Abdul-Alaziz Mohammed
Lecturer/Emergency Medicine Dept.
Mosul College of Medicine
Mosul University
Mosul/Iraq

Basim Mohammed
Asst. lecturer/computer center
Mosul university
Mosul/Iraq

Abstract:- A real time system embedded in mobile phone was designed in this work, called (Real Time Telemonitoring and Controlling System RTTCS) to telemonitor and control a patient's case in level two of telemedicine. The signals (ECG, Arterial Oxygen Saturation and Blood Pressure) were transferred out of the patient's monitoring equipments to NOKIA12 unit. Then they were send wirelessly through GPRS to be received by the mobile phone interpreted by the specialist physician who is far a way from the patient. By which the physician can evaluate the patient's case through parameters displaced on the mobile phone screen, so he can provide the necessary medical orders.

The suggested system consists of three units. The first is the NOKIA12 unit (T-Box N12 R) which contains an embedded real time program works as its operating system. That depends upon two principles multithreading and preemptive and uses a proposed dynamic scheduling algorithm called (RTT) with real time constraints to schedule the signals and to send them according to identified priorities to meet the deadline of signals. The second unit represents a web site which is the middle storage for the patient's data. The third unit is a mobile unit (mobile phone) which receives the coming signals from the patient monitor accordingly through the previously mentioned first and second units, then the physician can evaluate and diagnose the patient's case and order the necessary interventions.

The system was applied on many cases of abnormal cardiac rhythm cases, where it had been send successfully to a mobile phone in it's real time, and had been read by the physician where it was clear and reliable for the medical diagnosis.

Keywords: *Real Time, Operating System, Embedded, Telemedicine, Telemonitoring, GPRS, GSM, T-Box N12R, Multithreading.*

I. INTRODUCTION

The ongoing development in the field of wireless networks caused a huge progress in many scientific fields, especially, the field of telemedicine, which is concerned with the telemonitoring and controlling of medical cases. Since the use of mobile phones in the present time become very common, therefore, a lot of ideas in taking advantage of it for its distinguished characteristics, best of which is the mobility. Many programs and applications appeared to support the programming capabilities in this field, applying the real time principle in communication and data

transferring through the network reduced time of respond to the minimum [10].

The working of the mobile phone require a wireless network to control and manage the mobile units, this leads to the establishment wide wireless networks such as GSM and the development of many services such as GPRS for transferring data. The development of mobile applications entered many fields such as telemedicine, m-businesses, security, and a lot of other specialties because of the wide GSM spectrum compared with other types of networks [2].

Telemedicine is considered one of the most important mobile applications, which utilizes the activity of the international mobile communication network in the process of transferring data and it uses the available capabilities of the internet to demonstrate data to use them for the purpose of diagnosis and monitoring [4,10]. Telemedicine system considered one of the real time system as the evaluation of the task efficiency does not depend only on its accuracy but also on the time elapsed for its execution[9].

In telemedicine systems, the patients' signals are transferred at their real time to the health centre or the specialist doctor where it demonstrates an ultra short delay time in their transfer. Most of the telemedicine system are used in emergency cases which demand a rapid resuscitation, so for most of these systems, GPRS services is used because it offer a rapid speed for transfer in critical cases [9,7].

Telemedicine system consists of three parts [2], the first is the communication service provider, the second is the hospital service provider, and the third is the application service provider. Each one of them has its own special features and limits. Most of the applications in the field of telemedicine are basically depend on the principles of real time in data transmission to the final destination. So the data of the clinical cases are transferred to specialist doctor before the reach of the deadline. This is considered as one of the most important constraints of the real time systems [1]. Also the priority is considered one of the important limitations, which is determined by designer. All limitations which fall within the field of the applications enter a scheduling process for all application variables within specific algorithm [3].

Most of the real time systems are considered to be embedded systems, which mean that these systems are part of the physical environment. Generally the embedded

systems uses special purpose computers instead of general purpose computers, the design and the representation of these systems demands the following points for the signal [5, 8]:

- 1- Selection of the hardware and software and determine the cost for the real time systems.
- 2- Specified the design of the real time system and correctly determine its behavior.
- 3- Comprehend the programming language for implementing a real time system.
- 4- Increasing the reliability to the maximum and decrease the errors to the minimum for the real time system.
- 5- Testing the real time system.
- 6- Integration of all parts of the real time system.
- 7- Predictability response time.

II. RELATED WORKS

In 2007, Chen X., et al. where constructed a Holter monitor which is a mobile cardiac monitor connected directly to the patient and record the cardiac rhythm with the ability of detecting abnormal cases and recording them. Then after a while these cases are transferred to the mobile phone which should contain operating system (windows) with processing speed not less than 300 MHZ. The cases data are demonstrated on the mobile screen where interpreted by the doctor [6].

Cebrian A. et al., used the principle of real time on the design of remote monitoring in the year 2007. This system consisted of a memory, a microcontroller, and a power supply, the patient signals are stored on the memory system and depending on special algorithm the real time principle of the signals are achieved. After the signals passed to the algorithm it was send by GPRS to the doctor's computer, so he follow up the patient case [2].

Yanzheng, et al. design a remote monitoring system in 2007 and it depends on the presence of internet service, Linux operating system and the use of the CGI programs. Here the signals of the patients ECG are sent to the internet and then transferred to a computer which demonstrate the patient case on its screen [11].

In 2008, Xin G., et al. designed a real time system for remote patient monitoring. Where the design of the system needed the presence of a device connected to the patient's ECG device. This device is responsible of the transfer of the patient signals through the GPRS service and sending them to the internet through (UDP) protocol, then these data are sent to the emergency centre to interpretate them. This system has the ability to determine the location (GPS) [12].

III. CONTRIBUTIONS

The design of this system depends on the real time concepts, this system connect to the GSM network and internet. An embedded program has been designed which uses a real time algorithm to schedule special patient's signals. We choose three signals for the patient (ECG signal, Blood pressure signal and Oxygen saturation signal). These signals have been scheduled, transferred to GSM network, and then to the mobile phone. The design of the system achieves the following:

- 1- Design an embedded real time program work and operate as operating system for the Nokia12 unit, this program manages and operate the Nokia12 unit in most efficiency and precision. Multithreading has been used in the design of this program.
- 2- Design a proposed real time algorithm in the embedded program for Nokia12 unit called Real Time Telemonitoring (RTT) algorithm. This algorithm schedules the patient's signals and then sending them in their real time. The signals deadline and criticality are considered in determining sending priority.
- 3- Design Telemonitoring system in level2 of telemedicine for monitoring a patient's case especially if the patient is in a far away region and in emergency case. Using the GPRS as a carrier of the data makes the system faster in transferring them.
- 4- Using mobile phone from the specialist doctor for telemonitoring and telecontrolling makes the system more progress than other system because of the mobility feature, and the other feature is to change the location from place to another place in any time. All operations of the system depend on the Multithreading concept.

IV. SYSTEM OUTLINES

The designed real time system consists of three unit figure (1). These units cooperate with each other in streaming and synchronized manner. The system units are:

- 1- First Unit (T-Box N12R(Nokia12 Unit)).
- 2- Second Unit (Web Site).
- 3- Third Unit (Mobile Station).

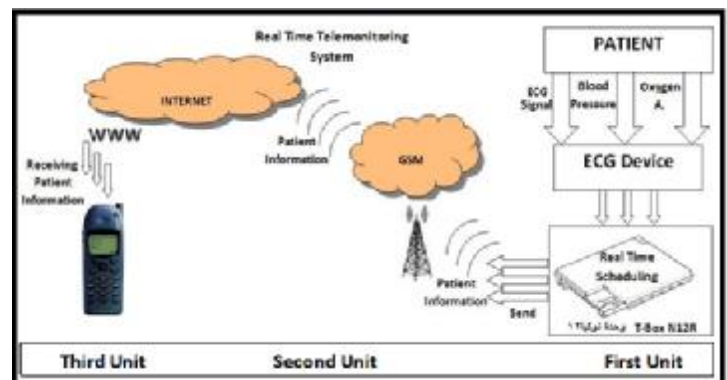
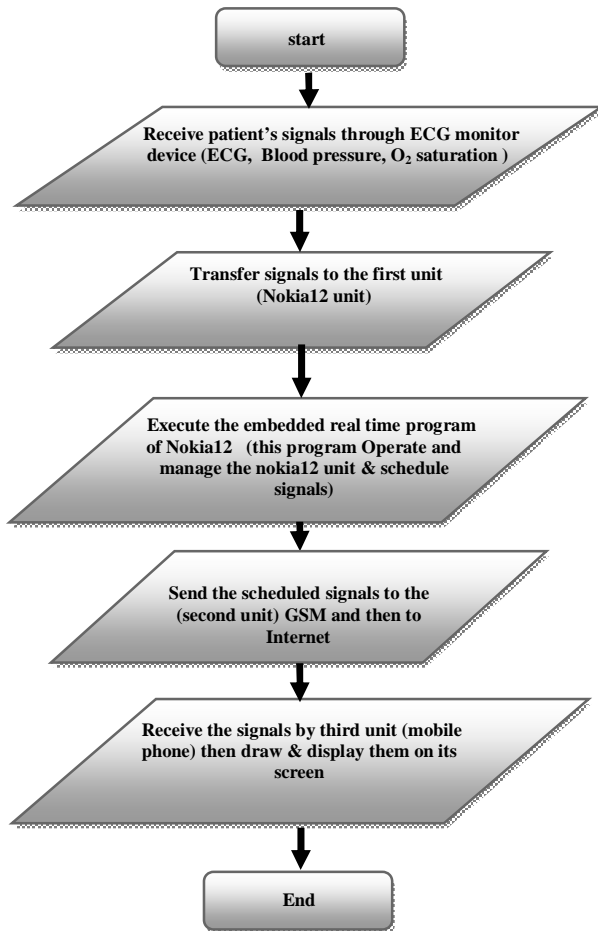


Figure (1) General System Design

The first unit of the system (Nokia12 unit) receives the data from the ECG monitor device. The embedded program (the designed operating system) for the Nokia12 unit is responsible for managing this unit and scheduling the patient's signal through a proposed Real Time Telemonitoring algorithm (RTT algorithm). When scheduling is completed, this unit sends the signals depending on criticality and specific priorities through GSM.

When the signals reached the second unit (web Site), the CGI programs written will save the signals in a file work as buffer then send it to the third unit (Mobile Unit) of the system. Patient's state will displayed on the screen of the mobile phone, the flowchart (1) show the flow of system work.



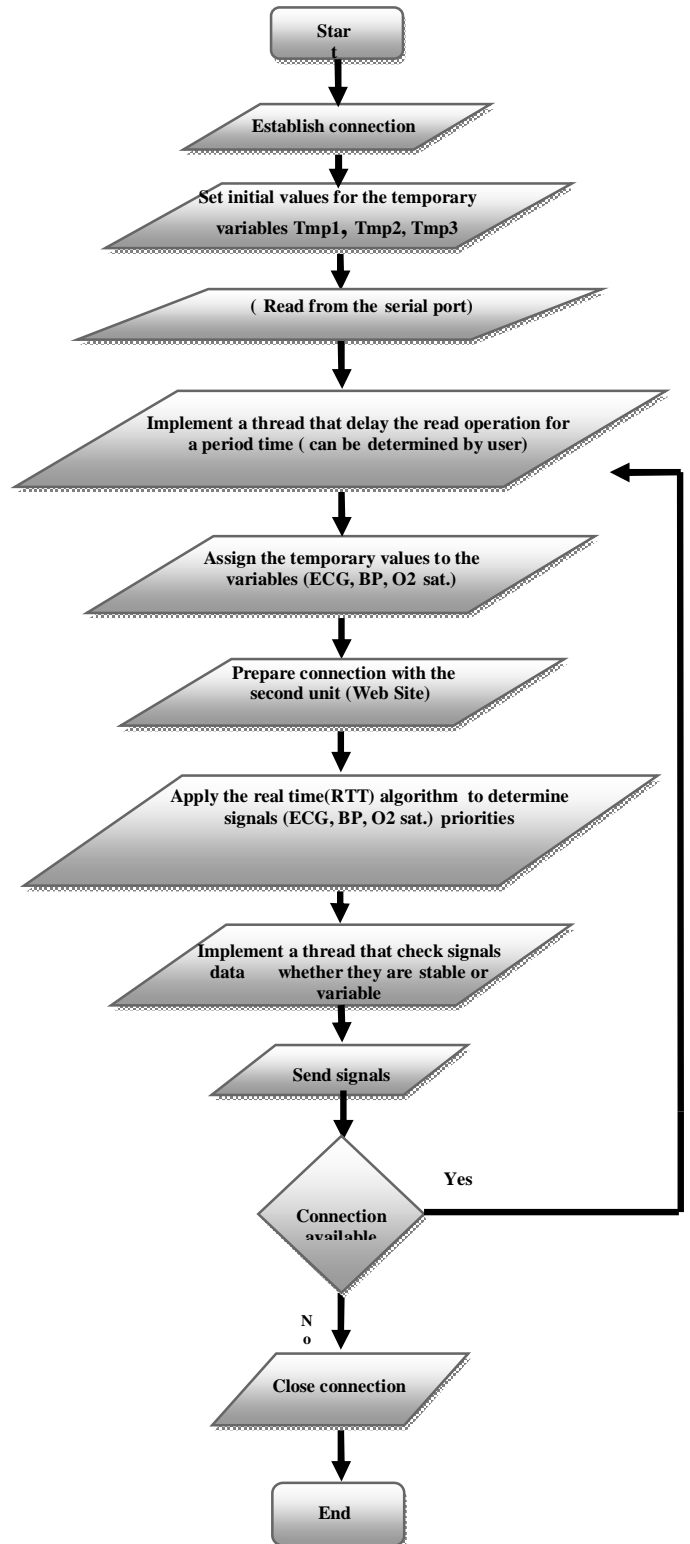
Flowchart (1): Real time system work flow

• First Unit: T-Box N12R (Nokia 12 Unit)

The embedded real time program of Nokia12 unit is shown in flowchart (2) which represents the embedded program of Nokia12 unit, this program depends basically on real time concepts to determine signals sending priorities, the program performs the following tasks:-

- 1- The embedded program of Nokia12 unit works as an operating system for this unit, since it transfers the data through the ports of this unit and scheduled the signals without user's interference.
- 2- Sending signals to the second unit is done in a real time manner depending on special algorithm (RTT algorithm). This algorithm schedules three signals (ECG, Blood pressure, Oxygen saturation) depending on criticality and priorities and then sends the signal that posses the highest priority first. The proposed algorithm has a sequential execution as below:-
 - a) Set initial priorities for the signals (ECG Signals, Blood Pressure, Oxygen Approximation).
 - b) Compute the execution time (cost) for signals.
 - c) Assure if any critical case in signals.
 - d) Make a comparison to candidate one of the patient's signals, re-compute the priorities again, the setting of the priorities is depending on the dead line and if critical case occurs.

- e) Send the signals depending on their priorities.
- f) Return to step a.
- 3- Establish connection between Nokia12 unit and the internet throw HTTP connection commands on the internet and GPRS.
- 4- Achieve continues loop for the data transfer.



Flowchart (2): Embedded real time program of Nokia12 unit.

This embedding program manages the Nokia12 unit, the program basically depend on the real time constraints for determining priorities of the signals used in this system. The program depends on many ports that included in Nokia12 unit. These ports are directly connected to the ECG monitor device which will send signals (ECG, Blood pressure, O₂ sat.) to Nokia12 unit.

At the beginning of the program, there is a process of establishing a connection to the GSM network and then to the GPRS in order to connect the device to the internet. Temporary variables (Temp1, Temp2, and Temp3) will set to initial values. These represent system variables that contain values from the patient by the ECG monitor. Reading these values is considered as periodic tasks with specified period time.

The real time embedded program depends on Multithreading technique through the execution. Two threads have been defined in the program, the first one for reading the system variables, and the second for checking the values of variables (signals) and sending them to GSM network. The benefit of multithreading concept is to coordinate the execution among more than one operation. It is possible to do reading process and at the same time checking values and sending them. So they can be executed in parallel manner.

Reading signals task period time is determined in the embedding program by 10 milliseconds. There is a possibility to change this time by the user of the third unit (mobile phone).

The proposed real time algorithm (RTT) in the embedding program inside the Nokia12 unit start executing for scheduling system variables and sending them depending on their priorities. Sending system signals does not happen until there is a change in the values of system variables. This is due to keep economical effectiveness of the system low and to reduce communication cost. At this time a connection activity is changed, if there is no connection, the system will reset connection variables. If there is then another read operation is done and etc.

◇ Building Connection with Nokia12 Unit

When the connection is established, special variables (is, os, c) are being defined that responsible for flowing data in and out of Nokia12 unit. This unit is connected to the server (Web Site) which is given this address (<http://rtt.freehostia.com>). The web site works as a temporary storage for the system data. This web site contains three web server programs (CGI programs) written in Perl language as a connection protocols between system units, POST method are proposed to manage the connection.

◇ Real Time Algorithm (RTT)

After determine the variables of the algorithm (the values of the system signals), the algorithm starts to determine signals priorities. The algorithm gives initial priorities for the system signals from highest to lowest (ECG_p, BP_p, OA_p), and then compute the execution time for the signals. Signals' deadlines are computed as follows, if T_{ij} represent state j from the task T_i, the absolute deadline is:

$$\text{Deadline } (d_{ij}) = \hat{O}_i + ((j-1) * P_i) + D_i$$

While (\hat{O}_i) represent the release time for the first period, (P_i) represent period and (D_i) represent the maximum response time (relative deadline).

◇ Priorities with Critical Cases

When a critical case is occurred, the program checks signals' values. If the ECG signal is the critical case, then it will get the first highest priority. Between the other two signals (Blood pressure, Oxygen sat.), the deadline is computed. The signal with the smallest deadline will get the second highest priority and the other will get the third priority. The same procedure will be executed if a critical case is occurred with the blood pressure or oxygen saturation values.

◇ Priorities without Critical Cases

In this case, the priorities of the signals (ECG, Blood pressure, and O₂) are determined depending on their deadlines. The signal with smallest deadline value will get higher priority. So after computing signals' deadlines, the priority of the ECG, Blood pressure, and oxygen saturation are 1, 2, and 3 respectively.

◇ Checking Data Changes

This part in the embedding program represents a thread, this thread check variables values, if there is a change in the values, the sending operation occurred else no sending is occurred. Since the sending operation has a cost according to the communication company, so applying this checking make the system cost effective.

◇ Closing connection

If the connection is interrupted or any exception occurred, the connection will be closed. This is done by putting value Null in the connection, sending and receiving variables. Then a call to special classes for the closing operation like (pause(), Destroy(), and Stop()) is done.

● Second Unit

After sending patient's signals from nokia12 unit to internet in a real time manner, there is a need to the server to act as temporary store for the signals data. This store represents an internet site called (<http://rtt.freehostia.com>) which receives the data and record them in text file named (rtt.txt). Then the user (doctor) could open this open this site and monitor the patient. Three CGI programs have been written to coordinate the three units of this system figure (2).

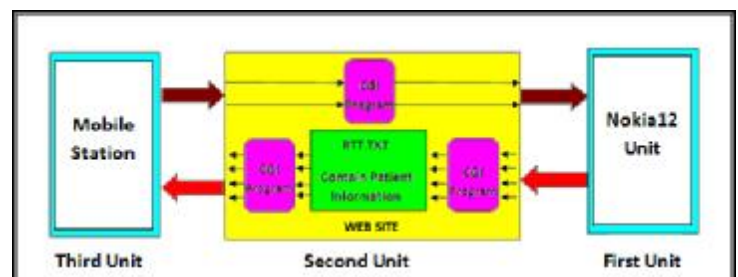


Figure (2): The General Structure Of The Second Unit

◇ Services programs

1-First web program (Tele1)

It is a CGI program written in Perl language to act as a protocol which coordinates the work between Nokia12 unit and the web site. This program receives the scheduled data from Nokia12 unit and stores them in a buffer then creates

a text file called (rtt.txt) where the data is transferred from the buffer.

2- Second web program (Tele2)

It is a CGI program works as a protocol to coordinate the work between the mobile phone and the web site. It is used to read the scheduled data stored in (rtt.txt) file and send them to the mobile phone to be displayed on its screen.

3- Third web program (Tele3)

This program is a protocol to coordinate the work between Nokia12 and Mobile Phone. By this program, the real time system can change period time between two readings of the patient's signals. This changing is done by the specialist doctor for the received signals to be sent to the second unit (web site). The period time value is saved in text file (period.txt). The data of this file will be sent to the real time program in the first unit (Nokia12 unit).

Table (1) Authorized Users Data Base

Phone No.	Tri Name	E-Mail	Retype password	Password	Username
Digits	Char(30)	Char(30)	Digits	Digits	Char
-	-	-	-	-	-

4- Site Management Program

This program is used to manage and design the internet site. It creates a special database for the authorized users' information (doctors) as in table (1). This was done by using the languages (Javascript, CSS, PHP, and SQL).

•Third Unit (Mobile phone)

This unit is the mobile phone that receives the data from the second unit (web Site) and displaying the result (Data) on the mobile screen. The specialist doctor could diagnose the patient's case by this information. The GPRS service must be available on the mobile phone, and the access point (AP) must be determined according to the communication company. Since transferring the data from the web site is done by connecting the mobile phone with the internet. The program of the third unit was written by J2ME language, which suitable for programming the smarty devices. The main function of this program is receiving data from the web site by server programs, processing this data and then drawing it on the mobile screen (algorithm 1).

Algorithm 1 : Mobile phone program

- 1- Begin
- 2- Create GUI
- 3- Establish connection to gather data to mobile phone.
- 4- Download the text file that contain the patient information (Values of signals).
- 5- Process the text file.
- 6- Draw the signals and show the alerts of the critical cases.
- 7- If exception occurred then go to step 4
Else wait to complete another period

- 8- If connection exception occurred then go to step 3
Else close connection
- 9- End

◇ Demonstrating Monitoring Signals

In the mobile phone, a program was designed to draw monitoring signals based on the data received from the web site. Special equations are prepared for drawing these signals. These equations are produced after many trials and studies of the ECG signals data. As the final ECG graph contain many waves (P-Q-R-S-T), each wave contribute to the built up of the final ECG signal. On which the diagnosis of the clinical condition is based.

Testing operation is began by testing the values S1 (Blood pressure), S2 (oxygen saturation), and ECG. If one of them is outside the range of the alarm which detects the emergency case, then the alarm will start to function. The alarms are set in this system as follows

- Oxygen less than 90%.
- Systolic blood pressure less than 100 and more than 180.
- Abnormal ECG signals.

In both cases (Critical and Normal) the program continues to draw the ECG signal, and demonstrate the values of S1, S2.

Two main concepts where depended upon in programming of the system, the preemption and multithreading. On downloading the signals from the web site, the ECG signals where drawn at the same time if any critical signal is detected, then the preemption concept is utilized. When there are two threads in the execution, the first thread is responsible for downloading the file, while the other for drawing the signals. The two threads are worked in parallel manner.

The algorithm of drawing ECG signals on the mobile screen is as follows:

- 1- Initiate special equation for drawing ECG.
- 2- Check values of signals variables (ECG, S1, S2)
- 3- If signals are not critical then
Draw ECG signal & display S1 & S2 (Implemented as thread)
Else activate critical cases alarms
- 4- End.

V. SYSTEM IMPLEMENTATION

Before mentioning how the system works, we will list the capabilities and characteristics of the system:

- 1- The specialist doctor has the ability to capture any case and store it in the mobile phone and create a document file for the patients.
- 2- The specialist has the ability to determine period time between two signals read.
- 3- The specialist doctor has the ability to determine signals priorities from highest to lowest.
- 4- The specialist has the ability to transfer from location to another, because of the large scale of GSM.

- 5- Nokia12 unit has the possibility to be in any international location, but must be provided by GSM network.
- 6- The simplicity of using the system by the specialist doctor.
- 7- The specialist has the ability and authorization to monitor the patient's data signals through web site.

◇ Web Site execution

The web site represent the second unit from two sides, the first one the internal function, the second is the interface that support the specialist doctor to see patient's information.

◇ Web Site Internal Function

The internal function of the site includes the operation of sending and receiving data from the first unit.

◇ Web Site as interface

The specialist doctor can see the patient's information which received from the first unit (Nokia12 unit) of the system, the web site address is (<http://rtt.freehostia.com>), figure(3) shows the general form of the site. The site is protected from unauthorized intruders. Adding another user is possible by giving him an authorization to enter the site by recording some information about him and give him username and password. After completing the recording of the new user's information, he can visit this site anytime. Patient's information stored as a text file, this file can be load and open with Microsoft Excel. The graph of the ECG signal can be seen but without grid. The first execution of the mobile phone program has been executed and tested by the emulator program called (WirelessToolkit), then embedding it to the mobile phone.



Figure (3) General site form

Figure (4) shows the execution of the system on the WirelessToolkit. The screen is partitioned to equivalents squares size of the values for ECG drawing purpose. The Blood pressure and Oxygen saturation signals are shown as numbers. Figure (5) shows commands system.

- 1-Operate:** this command is used for operating the system, establishing connection and receiving data from the first unit as shown in figure (6), this command change to "Stop" when the system is start working. When pressing "stop" command, the system stopped and the connection is closed.

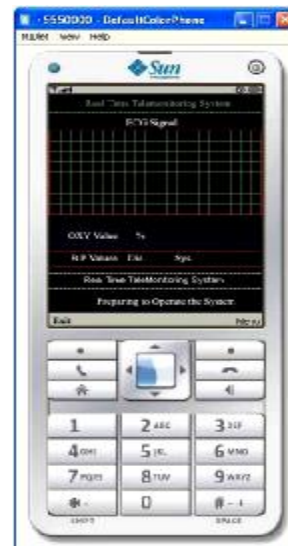


Figure (4) : General system form

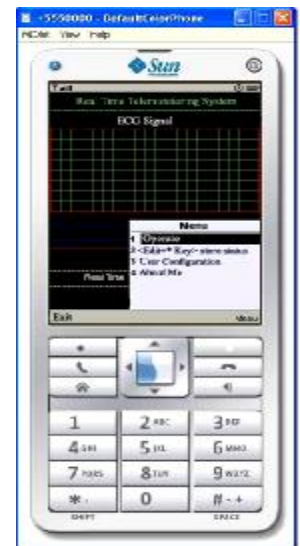


Figure (5) : Menu choices

- 2- <Edit Key + * Key> store status:** this command for capturing any important patient's case. This command depends on the application called (ScrnShot).

- 1- User Configuration:** by this command, the specialist doctor (mobile user) could do some configuration for the system. He can Change patient's signals priorities and period time between two signals reading. So this represent a protocol between mobile phone and Nokai12 unit (see figure 7).

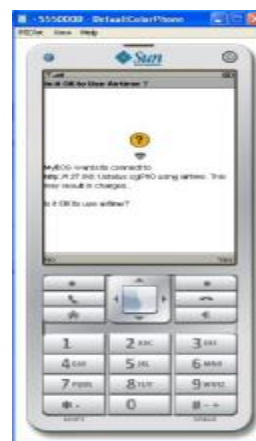


Figure (6): Connect to Internet

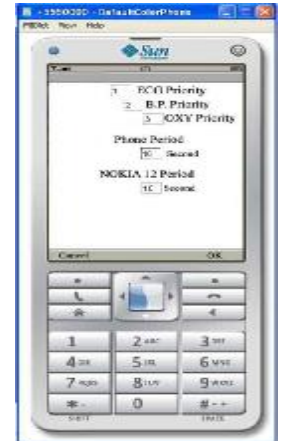


Figure (7) : User Configurations

VI. SYSTEM TESTING

The implementation and testing of this system was done on real cases data. Some of them may need urgent interventions. The system proved to be able to send the needed data specially these associated with the cardiac monitor. The data were uploaded to the Nokai12 unit and then they were sent to the mobile phone, which they demonstrated on its screen. These data enable the specialist doctor to diagnose the case depending on the data show process which occurs continuously on the mobile screen. The system was implemented on different cardiac rhythm cases, three of them are demonstrated in this paper. One of them is

the normal cardiac rhythm and the other two are abnormal rhythm cases.

Case 1:

A normal cardiac rhythm is demonstrated. The mobile screen also shows normal Blood pressure and Oxygen saturation figure (8).

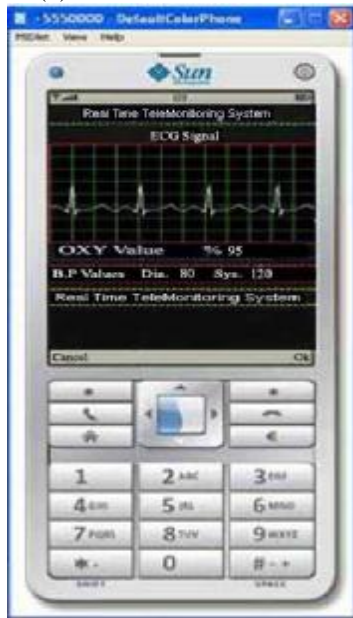


Figure (8): Normal Cardiac rhythm

Case 2:

Atrial fibrillation is shown, which is characterized by completely irregular cardiac ventricular electrical impulses (QRS Complexes). Absent atrial impulses (P-waves), which are replaced by Fibrillatory waves. The Blood pressure and Oxygen saturation also are seen on the mobile screen figure (9).



Figure (9): Atrial Fibrillation

Case 3:

Ventricular tachycardia is seen; in this case the excitation comes directly from the ventricles rather than from the atria to the ventricles as in normal situation. The heart rate will be more than normal (150-250) beat/min.), but they

are still regular and the ventricular wave impulses (QRS Complexes) are wider than normal. The atrial impulses are completely absent. Here the Blood pressure is seen to be below normal (Hypotension) and the Oxygen saturation is also low figure (10).

Upon testing this system the specialist doctor was able to diagnose the cases within the appropriate time limits for the intervention and treatment. This was supported by the sending patient's signals in real time manner.



Figure (10): Ventricular Tachycardia

◇ Computing signals' deadlines in Nokai12

The deadline for ECG, Blood pressure and Oxygen sat. Signals have computed depending on the values of (release Time, Execution Time). The release time for all signals in the first period is equal to Zero. The execution time is computed by the Nokai12 unit emulator. The deadline equation is:

$$\text{Deadline}(d_{ij}) = \hat{O}_i + ((j-1) * P_i) + D_i$$

The dead line values are the factor of the scheduling signals and set the priorities. Table (2) shows the deadline values for the first reading (first period) in Nokia12 unit for the three signals of the third case that are called (Ventricular Tachycardia).

Table (2): Deadline values for Ventricular Tachycardia case.

	ECG Signal	Oxygen Sat.	Blood Pressure
Release Time (ms)	0	0	0
Execution Time (ms)	9.3	10.5	13.6
DeadLine (ms)	9.3	10.5	13.6

VII. DISCUSSION AND CONCLUSIONS

In this work, a design and implementation of an embedded real time telemonitoring and controlling system was done for the mobile phone. This system was used for monitoring and controlling medical cases in level 2 telemedicine. In this work monitoring of the electrocardiogram, Blood pressure, and Oxygen saturation signals was done. The suggested real time system was designed, implemented, and tested in the lab where all the

system procedures and signal transformation were assured to reach at their real time to the mobile phone. In the following are some medical and software conclusions:

◇ Medical Conclusions:

(Level 2 Telemedicine): was selected because it has the advantage of transferring the information to the far away situated specialist doctor at their real time. This is of special importance in emergency cases where the delay may lead to disastrous events.

(Web Site): Designing a special web site for the system gives the following advantages:

- a- It gives the ability for the specialist to monitor and follow up the patient through the internet on PC where bigger screen helps him in better interpretation of the signals.
- b- Storing all patient data to revise them later for medico-legal or teaching purposes.

(Configuration): The doctor in this system has the ability of determination signals priorities according to their importance for the medical diagnosis. Also he can determine the received signals periods in Nokia12 unit and reading signals periods from the internet to the mobile phone, so it enables him to monitor every change on ECG signal. He has also the ability of freezing and capturing events for further analysis.

(LEAD II ECG): Selection of lead II ECG in the monitoring of the cardiac rhythm was done because it is the most usual and frequently used lead in continuous cardiac rhythm monitoring.

◇ Computing Conclusions

(Real Time): the using of real time concept is compatible with the international needs in all systems in the medical scope, especially the telemedicine that the main factor of the response in certain time.

(GSM & GPRS): the main features of this network is the wide coverage area, this network gives any features distinguish from another wireless networks, especially the system that needs speed and precision in transferring data. The GPRS service supports the cost effectiveness. If there is no data transferring, no cost is paid.

(Traffic Lost and reduction): Transferring ECG signals of abnormal cases through GSM network to the mobile phone and the web site leads to reduce the cost and traffic jam on the network.

(Multithreading): The use of multithreading technique in the operating system programming of Nokia12 unit and mobile program leads to achieve the coordination and the synchronization of program threads parallel execution and reduce the execution time. This is necessary in embedded real time systems. Also this property allows the program to be more scalable.

(Preemption): The use of this property gives the ability to preempt low priority task by higher priority task (most critical).

VIII. FUTURE WORKS

There are many suggestions for the future work:-

- 1- Connecting more than one Nokia12 unit in many regions to form what is known as (Distributed Real Time

System), in this case the system works in wider level (more than one city or country).

- 2- Developing an embedded program in Nokia12 unit to transfer multimedia data when these services are available in the GSM network. This will lead to make an advanced system which has the ability to transfer images or videos for patient at their real time e.g. Real Time Echocardiography.
- 3- Studying the ability of connecting Nokia12 unit to GPS unit for the determination of the exact patient's location.
- 4- Connecting the system to special devices which can give medications. The system should be provided by control panel for these medications. This property will transfer the system from level 2 to level3 telemedicine.

IX. References

- [1] Albazaz D., "Design and implementation of real time software for multitask system", ph.D., College of computer and Mathematical sciences, University of Mosul, Iraq, 2004.
- [2] Cebrian A., Guillen J., and Millet J., "Design of a Prototype for dynamic electrocardiography monitoring using GSM technology: GSM HOLTER ", Proceedings of the 23rd Annual EMBS International Conference, IEEE 2001 October 25-28, Istanbul ,Turkey, PP: 3956-3959, 2001.
- [3] Jeffay K., Becker D., and Bennett D., "The design, implementation, and use of a sporadic Tasking model", university of north carolina at chapel Hill Department of computer sciences, Chapel Hill, NC 27599-3175 USA, 1994.
- [4] Joon S., Hang L., Ho S., and Shuji S., "Telemedicine System Using a High-Speed Network: Past, Present, and Future", Gut and Liver, Vol. 3, No. 4, pp. 247-251, December 2009.
- [5] Laplante, A., "REAL-TIME SYSTEMS DESIGN AND ANALYSIS", 3rd Ed., Published simultaneously in Canada, 2004,.
- [6] Chen, X, Lim E., and Kyaw T., "Cellular Phone Based Online ECG Processing for Ambulatory and Continuous Detection", IEEE, Institute of Infocomm Research, SG, Singapore, pages 653-656, 2007.
- [7] Halima Ali Abd Elazeez, "Distance Learning & Telemedicine & Participation in Medical Education", College of Medicine Conference (CMCI), Plistain University, 2005.
- [8] Prasad K., "Embedded/Real time systems: concepts, Design and programming", Dreamtech press, 19-A, Ansari Road, Daryaganj, NewDelhi, 2005.
- [9] Qiang Z., and Mingshi W., "A Wireless PDA-based Electrocardiogram transmission System for Telemedicine", IEEE, Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, September 1-4, pages 3807-3809, 2005.
- [10] Sanaullah C., Humaun K., Kazi A., and Kyung-Sup K., "A Telecommunication Network Architecture for Telemedicine in Bangladesh and Its Applicability", International Journal of Digital Content Technology

and its Applications, Volume 3, Number 3, September 2009.

- [11] Yanzheng, L., Shuicai W., Jia L., and Yanping B., "The ECG Tele-monitor Based on Embedded Web Server", IEEE, Biomedical Engineering Center, Beijing University of Technology, Beijing, China, pages 752-755, 2007.
- [12] Xin G., Dakun L., Xiaomei W., and Zuxiang F., "A Real Time Continuous ECG Transmitting Method through GPRS with Low Power Consumption", IEEE, Department of Electronic Engineering, Fudan University, Shanghai, China, Pages 556-559, 2008.



Dr. Dhuha Basheer Abdullah Albazaz /Asst. Prof / computers Sciences Dept. / College of Computers and Mathematics / University of Mosul. She has a Ph.D. degree in Computer Sciences since 2004. Specific Specialist in Computer Architecture and Operating System.

Supervised many Master degree students in operating System, computer architecture, dataflow machines, mobile computing, real time, distributed databases. She has three Phd. Students in FPGA field, distributed real time systems, and Linux clustering. She also leads and teaches modules at both BSc, MSc, and Phd. levels in computer science. Also she teaches many subjects for Ph.D. and master students.



Dr. Muddather Abdul Aziz Mohammed/Lecturer / emergency medicine/ Mosul College of Medicine /University of Mosul. Honored the degree of higher specialization and the certificate of Jordanian medical council in accident and emergency medicine in 2002 ,

supervisor of Mosul center of Arab Council of Health Specialization in emergency medicine since 2007. Beside under and postgraduate teaching he supervised and organize many courses in emergency medicine and resuscitation both in IRAQ and JORDAN .

Automating the fault tolerance process in Grid Environment

Inderpreet Chopra

Research Scholar
Thapar University Computer Science Department
Patiala, India
inderpreet@thapar.edu

Maninder Singh

Associate Professor
Thapar University Computer Science Department
Patiala, India
msingh@thapar.edu

Abstract:

As Grid encourages the dynamic addition of resources that are not likely to be benefited from the manual management techniques as these are time-consuming, unsecure and more prone to errors. A new paradigm for self-management is pervading over the old manual system to begin the next generation of computing. In this paper we have discussed the different approaches for self-healing the current grid middleware use, and after analyzing these we have proposed the new approach, Self-healing Management Unit, SMU that will provide the automated way of dealing with failures.

Keywords: SMU, heartbeat

1. Introduction

In recent years Grid, which facilitates the sharing and integration of large scale, heterogeneous resources, has been widely recognized as the future framework of distributed computing [1, 2]. However, the increasing complexity of Grid services and systems demands correspondingly larger human effort for system configuration and performance management, which are mainly done in a manual style today, making it time-consuming, error-prone and even unmanageable for human administrators.

Autonomic computing [4, 5] presented and advocated by IBM, suggests a desirable solution to this problem. The vision of autonomic computing is to design and build computing systems that possess inherent self-managing capabilities [6]. In this paper a Self-healing model- SMU (Self-healing management unit) has been described which by autonomic computing has targeted to improve the level of automation and self-management capabilities to a far greater extent than it is today in Grid Computing systems. The SMU aims to:

- keep track over the jobs submission and execution
- recover the lost jobs
- administrate complexity of grid
- check the efficiency for resource discovery
- keep all the services running.

2. Self-Healing Mechanisms

Self-healing [7] is the ability of a system to recover from faults that might cause some parts of it to malfunction. For a system to be self-healing, it must be able to recover from a

failed component by first detecting and isolating the failed component, taking it off line, fixing and reintroducing the fixed or replacement component into service without any apparent overall disruption. A self-healing system also needs to predict problems and take actions to prevent the failure from having an impact on applications. The self-healing objective must be to minimize all outages in order to keep the system up and available at all times.

There can be many different reasons that can lead to the fault occurrence in grids. Some of the reasons we are able to find are as follows:

- *Hardware Faults:* Hardware failures take place due to faulty hardware components such as CPU, memory, and storage devices [8].
- *Application and Operating System Faults:* Application and operating system failures occur due to application or operating system specific faults like memory leakage, deadlocks, inefficient resource management etc.
- *Network Faults:* In a grid, computing resources are connected over multiple and different types of distributed networks. As a result, physical damage or operational faults in the network are more likely [9]. The network may exhibit significant packet loss or packet corruption. Moreover,

individual nodes in the network or the whole network may go down.

- *Software Faults:* There are several high resource intensive applications running on grid to do particular tasks. Several software failures like the unhandled exception; unexpected input etc. can take place while running this software application.

In addition to ad-hoc mechanisms – based on users complaints and log files analysis – grid users have used automatic ways to deal with failures in their Grid Environment. To achieve the automatic ways to deal with failures, various fault tolerance mechanisms are there. Some of these self-healing mechanisms are:

Application-dependent: Grids are increasingly used for applications requiring high levels of performance and reliability, the ability to tolerate failures while effectively exploiting the resources in scalable and transparent manner must be integral part of grid computing resource management systems. Support for the development of fault-tolerant applications has been identified as one of the major technical challenges to address for the successful deployment of computational grids [10]. To date, there has been limited support for application-level fault tolerance in computational grids. Support has consisted mainly of failure detection services or fault-tolerance capabilities in specialized grid toolkits. Neither solution is satisfactory in

the long run. The former places the burden of incorporating fault-tolerance techniques into the hands of application programmers, while the latter only works for specialized applications. Even in cases where fault-tolerance techniques have been integrated into programming tools, these solutions have generally been point solutions, i.e., tool developers have started from scratch in implementing their solution and have not shared, nor reused, any fault tolerance code. A better way is to use the compositional approach in which fault-tolerance experts write algorithms and encapsulate them into reusable code artifacts, or modules.

Monitoring Systems: In this a fault monitoring unit is attached with the grid. The base technique which most of the monitoring units follow is heartbeating technique. The heartbeating technique [11] is further classified into 3 types:

- *Centralized Heartbeating* - Sending heartbeats to a central member creates a hot spot, an instance of high asymptotic complexity.
- *Ring Based Heartbeating* - along a virtual ring suffers from unpredictable failure detection times when there are multiple failures, an instance of the perturbation effect.
- *All-to-all heartbeating* - sending heartbeats to all members, causes the message load in the network to grow quadratically with group size,

again an instance of high asymptotic complexity

Checkpointing-recovery: Checkpointing and rollback recovery provides an effective technique for tolerating transient resource failures, and for avoiding total loss of results. Checkpointing involves saving enough state information of an executing program on a stable storage so that, if required, the program can be re-executed starting from the state recorded in the checkpoints. Checkpointing distributed applications is more complicated than Checkpointing the ones which are not distributed. When an application is distributed, the Checkpointing algorithm not only has to capture the state of all individual processes, but it also has to capture the state of all the communication channels effectively. Checkpointing [12] is basically divided into 2 types:

- *Uncoordinated Checkpoint:* In this approach, each of the processes that are part of the system determines their local checkpoints individually. During restart, these checkpoints have to be searched in order to construct a consistent global checkpoint.
- *Coordinated Checkpoint:* In this approach, the Checkpointing is orchestrated such that the set of individual checkpoints always results in a consistent global checkpoint. This minimizes the storage overhead, since only a single global checkpoint needs to be maintained on

stable storage. Algorithms used in this approach are blocking and nonblocking.

Fault Tolerant Scheduling: With the momentum gaining for the grid computing systems, the issue of deploying support for integrated scheduling and fault-tolerant approaches becomes a paramount importance [13]. For this most of the fault tolerant scheduling algorithms are using the coupling of scheduling policies with the job replication schemes such that jobs are efficiently and reliably executed. Scheduling policies are further classified on basis of time sharing and space sharing.

3. SMU Approach (Self-Healing Management Unit)

Our approach is to provide efficient self-healing functionality in grids. Most of the existing approaches follow the heartbeat technique to find the working nodes. But in this new approach another way of handling the failures has been proposed. In SMU there is one central node for each cluster called Managing Node (MN) (figure 1). MN is responsible for all the job submission and the job execution of applications submitted to grid. Apart from this, MN also keeps on monitoring the status of the job at different time intervals, to make sure that the job is executing well.

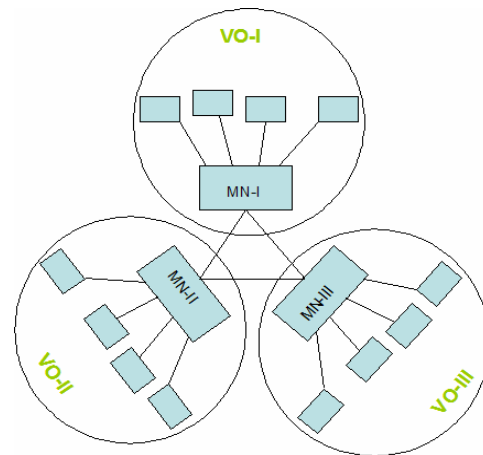


Figure 1: Grid Network

MN comprises of (figure:2) Application Receiving Unit (ARU), Node Receiving Unit (NRU), Processing Unit (PU) and database. ARU deals with the reception of request for the job to be executed, NRU helps to gather the status of the execution nodes attached to the MN. NRU is responsible for monitoring both the requests and responses in between the MN and the execution nodes. PU takes input from ARU, NRU and DB to decide to which execution node the job has to be assigned. PU uses the following algorithm to decide the execution of the application.

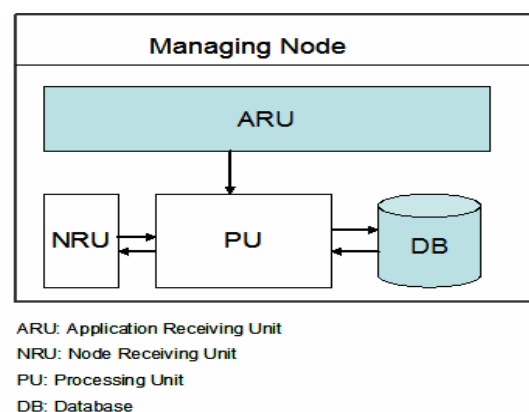


Figure 2: SMU Managing Node

Algorithm:

1. Check for the Current status for the execution node that is going to handle the job.
CHK 'STATUS'
2. If the status is "OK", then assign Job to that node. Else, find another node. Go to step 1.
3. Once node is selected, to check the status of the job, PU asks NRU to keep on monitoring the node after certain interval of time. For optimizing this process, we follow following steps:
 - a. Maintain the records for the kind of the job and their expected execution time in the database. Based upon that, start the monitoring after that expected interval, if PU doesn't receive any input from NRU.
 - b. If the job type is new i.e. not having any information, save it in the history database after it finish execution. But for monitoring it, decide any time say X minutes, after which the monitoring starts.
4. Node self-failure detection option is also available for the node to self detects the failure and send the status to NRU. This failure alarm can be raised based upon the CPU usage, insufficient resources to execute the job, etc.

4. Results

To test SMU, we submit jobs and introduce the faults into the setup grid environment. For producing the faults, we shutdown few nodes when job is submitted to them, another way we used to push the faults is by overloading the nodes which are used to execute the job. SMU do the auto job resubmission when some erroneous happens during job execution. We have found that using SMU, we can prevent the large number of jobs to get lost. To test this, we have started submitting large number of jobs and gather the detail that how many jobs got lost. Figure 3, shows that the rate of jobs executed in normal scenario is less as compared to SMU system. SMU helps the system to overcome from faults thus helps to increase the job execution rate.

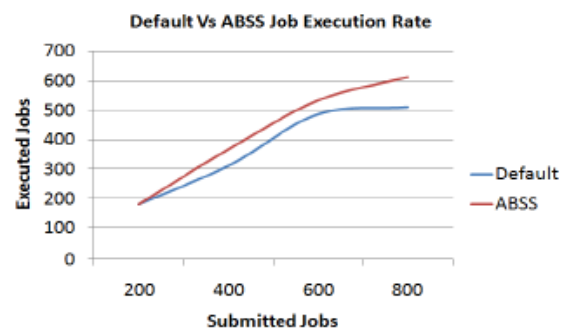


Figure 3: Default vs SMU execution rate

SMU also reduces the resources usage by optimizing the monitoring procedure (figure 4). If there are n number of nodes, and they keep on sending their alive signals at interval of t seconds, then the total time the managing nodes remains busy is $n*t$. But in the present approach, it reduces to large percentage as only those nodes are monitored that are needed for the execution of submitted job.



Figure 4: Optimized Monitoring reduces resource usage

It increases the scalability of resources. Now more resources can be attached with the single node. In normal heartbeat technique, it will be difficult to monitor nodes if the number of execution nodes increases. It also speed's up the execution of jobs whose entry is already there in the history table.

5. Conclusions:

As grid is a distributed and unreliable system involving heterogeneous resources located in

different geographical domain, for this case fault tolerant resource allocation services [14] have to be provided. In particular, when crashes occur, tasks have to be reallocated quickly and automatically, in a completely transparent way from the user's point of view. Different grid middleware uses different strategies to overcome faults. Like Globus [15] and Alchemi [16] uses the heartbeat to know about the current status of its execution nodes. On the other hand SUN N1GE6 [17] uses the user level and kernel level checkpointing strategies to restart the job from where it fails. Condor [18] provides system-level checkpoints of distributed applications but it not geared to high-performance parallel programs.

The proposed SMU approach is destined to increase the efficiency by reducing the unnecessary monitoring of nodes that are not going to participate in the execution of the job. At same time, information from the past execution of similar jobs is maintained, that determines the time interval after which monitoring of the job is started. As the execution nodes in the grids increase the SMU approach becomes all the more efficient.

References:

- [1] Ian Foster, Carl Kesselman, Jeffrey M. Nick, and Steven Tuecke. The physiology of the Grid. Global Grid Forum, June 2002.
- [2] Ian Foster, Carl Kesselman, and Steven Tuecke, "The anatomy of the Grid", Copyright 2003 John Wiley & Sons, Ltd.
- [3] Manish Parashar and Salim Hariri. Autonomic Computing: An Overview. Springer-Verlag Berlin Heidelberg, 2005.
- [4] Jeffrey O. Kephart and David M. Chess. "The Vision of Autonomic Computing", IEEE Computer, January 2003.
- [5] A. G. Ganek, T. A. Corbi, "The Dawning of the Autonomic Computing Era", IBM Systems Journal, .42 n.1, p.5-18, January 2003
- [6] Hang Guo¹, Ji Gao², Peiyu Zhu, Fan Zhang, "A Self-Organized Model of Agent-Enabling Autonomic Computing for Grid Environment", Proceedings of the 6th World Congress on Intelligent Control and Automation, June 21 - 23, 2006, Dalian, China
- [7] P. Horn. Autonomic computing: ibm's perspective on the state of information technology, October 2001. <http://www.research.ibm.com/autonomic/>.
- [8] Gerald Tesauro and David M. Chess and William E. Walsh and Rajarshi Das and et.al, "A Multi-Agent Systems Approach to Autonomic Computing", AAMAS'04 ACM (Jul 2004)
- [9] Luis Ferreira, Viktors Berstis and et.al, "Introduction to Grid Computing with Globus", http://www.liv.ac.uk/escience/beowulf/IBM_globus.pdf, IBM

- [10] Anh Nguyen-Tuong, "Integrating Fault-Tolerance Techniques in Grid Applications" PhD Thesis, University of Virginia, August 2000
- [11] Amit Jain and R.K. Shyamasundar, "Failure Detection and Membership Management in Grid Environments" Fifth IEEE/ACM International Workshop on Grid Computing (GRID'04) pp. 44-52
- [12] Sriram Krishnan, "An Architecture for Checkpointing and Migration of Distributed Components on the Grid" PhD Thesis, Department of Computer Science, Indiana University, November 2004
- [13] J.H. Abawajy, "Fault-Tolerant Scheduling Policy for Grid Computing Systems", IPDPS'04
- [14] Kamana Sigdel. Resource allocation in heterogeneous and dynamic networks. Master's thesis, Delft University of Technology, 2005.
- [15] Paul Stelling, Ian Foster, Carl Kesselman, Craig Lee, Gregor von Laszewski , "A Fault Detection Service for Wide Area Distributed Computations" Volume 2, Number 2, springer, pp. 117-128(12), 1999
- [16] Krishna Nadiminti, Akshay Luther, Rajkumar Buyya, "Alchemi: A .NETbased Enterprise Grid System and Framework" December 2005
- [17] Liang PENG, Lip Kian NG, "N1GE6 Checkpointing and Berkeley Lab Checkpoint/Restart" Dec 28, 2004
- [18] James Frey, Todd Tannenbaum, Miron Livny Ian Foster, Steven Tuecke, "Condor-G: A Computation Management Agent for Multi-Institutional Grids" 10th IEEE International Symposium on High Performance Distributed Computing (HPDC-10 '01)

A Computational Model for Bharata Natyam Choreography

Sangeeta Jadhav, S.S Dempo College of Commerce and Economics, Panaji, Goa India. sangeetajadhav@yahoo.com
Sasikumar, CDAC, Mumbai, India. the.little.sasi@gmail.com

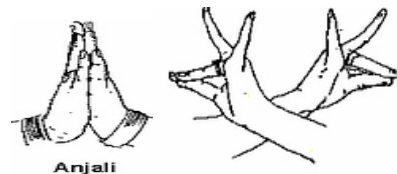
Abstract: Indian Classical dance has been the slowest to adopt technology and although choreography is mainly a domain of creativity, computers can help a lot to ease this especially for rhythmically oriented intricate footwork, *Nritta* and the complementary movements of other limbs of the body. The main objective of this paper is to identify and classify the various *angalakshanas* of Bharata Natyam. An *angalakshana* in Bharata Natyam refers to classification of elementary body movements like head, eye, neck, hands and leg movements. The research aims at a well organised classification of the *angalakshanas* along with identification of various constraints on their co- occurrences and sequencing. The constraints may be physical, aesthetic or preferential and would be identified through literature study and machine learning from databases of existing dance repositories and training videos. The work will be useful to enhance the learning/teaching of Bharata Natyam (and other dance forms) and also for animated choreography.

1. Introduction - To develop a computational model of Bharata Natyam covering *angalakshanas* and *talas* (a rhythmic cycle of beats with an ebb and flow of various types of intonations resounded on a percussive instrument). Any dance form can be conceptually decomposed into some constituent base movements involving specific body parts, and norms for combining them. The combinations are constrained by physical constraints, aesthetic constraints, preferential constraints, etc. Treatises like *Natyashastra* codify such constraints and rules in a human understandable form. This research attempts to model these in a computational framework, using practical experience and existing literature. This would help evolving better teaching programs, better understanding of the dance form, enhanced ability to compare one dance form to others in India and abroad, use ICT in composing and designing dance programmes, and so on.

The basic methodology proposed would be to study and tabulate the various *angalakshanas*, relate them to *tala* – a beat cycle – to study constraints and preferences, and develop a machine learning model which can generate/critic movement sequences. We propose to use artificial intelligence techniques such as neural networks, constraint programming, evolutionary programming and heuristic search in this process.

1.2 Structure of Bharata Natyam: Bharata Natyam is one of the most ancient of all the Indian Classical dance styles. Although there is no formal notation available for the dance form, it is a highly formalized dance as described in ancient texts. This dance is known for its grace, purity, tenderness and sculptural poses. Bharata Natyam is based on the precepts of among others - Bharata's *Natyashastra* and Nandikeshwara's *Abhinayadarpana* - to mention the most well known theorists of the dance. From the moment the dancer enters and until dancer leaves the stage, dancer has to regulate their movements to musical accompaniment in strict accordance with the rules of the dance form. There are two major aspects to this form of dance: *Nritta* (rhythmic dance movements) and *Nritya* (representational dance). Although there is no clear demarcation between the two this research will restrict itself to *Nritta* or pure dance movements. *Nritta* is divided into *Chari* - walking movement; *Karna* - balanced composition of hand and leg movements; *Angahara* - different combination of *Karna*; *Mandala* - postures with hand gestures. The body limbs are classified as: *Anga* (main body parts); *Pratyanga* (minor limbs); and *Upanga* (features of the face). *Pratyanga* and

Upanga are supposed to move in unison with *Anga*. *Anga lakshana* is a classification of elementary body movements which are as follows- Head movements-*Shirobheda*; Neck movements -*Greevabheda*; Eye movements -*Drushtibheda*; Leg movements -*Padabheda*; Hand movements - *Hastabheda*. Some of them are as follows:



Katoaka-vardhana

While performing, a dancer can use one *Angalakshana* like only the eye for example for a beat or use all the *Angalakshanas* e.g. eye, neck, head, legs, hands simultaneously for the same beat. Of course, not all combinations are physically possible. Some are aesthetically not recommended. Similarly, the sequence of *angalakshanas* from one beat to next also has such constraints governing them. This is the focus of the research we discuss here.

2. Related work in this area: Although many research papers can be found dedicated to Western dance styles like ballet and foxtrot, very few can be found for Indian Classical Dance. here seem to be few studies related to creating computerized choreography for any Indian classical dance. However the following papers and articles are related to Indian classical dance with emphasis on animation: Soumyadeep Paul, et al, [6] describes in their paper about an off-site coach who uses low band-width joint-motion data and can instruct students who are even geographically far away. This application does

not require the user to wear any kind of equipment. Sumant Narayan Pattanaik [7] in his article has described the work done at the National Centre for Software Technology, now CDAC in Mumbai on creating a computational modeling of human body and movement for animating Bharata Natyam dance.

2.1 Dance Notation

One of the problems involved in creating computational models for dance is that of notation. In the field of Western Dances 'Labanotation' a system using symbols to write down the movements of dancers, especially in a ballet has been developed particularly following the research conducted at the University of Frankfurt [9] and Ohio State University[9]. Using labanotation several software programs have been developed such as LabanWriter[15], CALABRAN[13], LabanPad[10&12], Limelight[11] and LifeForms[14]. However most of these programs are useful mainly either for notating and archiving dance or for animating dance. Annemette P Karpen, Bharata Natyam dancer and choreographer since 1977, presented a paper[16], "Labanotation for Indian Dance" at the 11th European conference for Modern South Asian studies, Amsterdam University, July 1990. In this she tried to solve the problem of notation for Bharata Natyam by using Labanotation a notation used for Western dance. However the problem was not completely resolved because of the many small movements of Bharata Natyam; hence she tried to adapt Sutton Shorthand Notation in which she added some extra signs for this purpose although this was not completely systematized.

Thus Labanotation is not adequate for Bharata Natyam and our proposed research focus however is totally different from the work mentioned above as it seeks to generate permutations and combinations of angalakshanas, for a given input beat, for which new algorithms may need to be developed. The idea of creating a new notation is not in the purview of this research at present.

3. Proposed Computational Model: The overall objective, as mentioned earlier, is to build a computational model of a dancer's body limb movements across time in the *Nritta* context. We started with a study and analysis to identify the various distinct movements possible in Bharata Natyam of various body parts such as eye, head, neck, hand, leg, etc. We have now identified a collection including:

- 9 Head movements(Shirobheda);
- 1 Neck movement (Greevabheda);
- 8 Eye movements(Drishtibheda);
- 39 Leg movements(Padabheda);
- Hand movements(Hastabheda):
 - ✓ 14 single hand(asamyukta)
 - ✓ 3 double hand (samyukta)

Typically at any point in time, a dancer can choose for each body part one of these possibilities. Where there is no explicit movement, the default position can be considered. Based on experience, observation and literature we have identified movement combinations which are either infeasible or unusual. This is currently within one time unit (one count/beat/taal).

For formalizing the model, we have adopted a notation system as follows. We use the Sanskrit name for the Angalakshanas as mentioned before; the first two letters of this name is used

(sometimes three or four in case of similar ones) to indicate a position of that part for e.g. code starting with SH indicate Samyukta Hasta which is double hand movement and GB indicate Greeva Bheda which is a neck movement and so on. The various positions of a limb are given as a two character code e.g. AH-pa indicates pataka from Asamyukta Hasta. Thus across time SH-an, SH-kat, SH-aw means double hand movements moving from one position to another position. This enables us to define a dancer's position at a particular point in time as a vector giving the position code for each limb. For example in following fig.



■ [SB-ad, DB-av, SHdol, PB-m-sw]

Shiro Bheda- Adhomukha
Dhrishti Bheda-Avalokitam
Samyukta Hasta-Dola
Pada Bheda- mandala-

swastika

This pose can be vectorised as follows:

[SB, GB, DB, SH, PB].

Across time, each element of the vector changes the code. A sequence of vector represents a dance movement in time duration. This structure will be used for further analysis and modeling.

The next concern addressed was movement feasibility. As mentioned earlier this could be physical constraint (sitting and standing at same time or bending back and front or moving neck side to side and looking straight at same time is not possible), dance theory base(hip and chest movements are not allowed), preferential etc.

A physical constraint within a vector has been identified. The restriction may arise more often not as a general restriction, but as applicable in certain positions e.g. if neck is in position 3 at time t1 then in (t1+1) position 1 is not feasible. Identifying such constraints along with the applicable context can be a hard task given the exponential number of combinations possible. Not that while we expect most constraints to be either at a particular time point or between two successive time points, it is possible to have constraints across more than two time units also. We do not consider this at present.

We propose to collect many instances of possible dance position vectors (henceforth dp vectors), have them labeled by human dance expert and use them to generate the knowledge base of such constraints. We have completed a simple vector generator and further refinement and labeling is in process.

These labeled instances will be used in a machine learning program to build the model. Inductive learning, evolutionary programming and neural network model are under consideration at present. While a model like inductive learning can produce human understandable knowledge, compared to models like Neural Network, other aspects such

as quantity and quality of data and nature of data need to be also considered in making a selection.

Among the current research challenges are formulations of an appropriate feature vector for learning algorithm, choice of learning algorithm, representation of context, identification of parameter outside the dp vector which may influence the choices, dealing with variability among human expert judgment etc.

The proposed tools to be studied viz. - trained neural networks; genetic algorithms and heuristic searches have been chosen for the study from among others because of the computational complexity of the problem. The utility of artificial neural network models lies in the fact that they can be used to infer a function from observations. This is particularly useful in applications where the complexity of the data or task makes the design of such a function by hand impractical. A simple set of *angalakshanas* or elementary body movements has following complexity: 9 Head movements; 4 Neck movements; 8 Eye movements; 39 Leg movements; Hand movements: 28 single hand and 23 double hand movements. For choreography of n beats for each of these the permutation and combination will result in $(7233408)^n$ movement possibilities! The resulting database for any beat is huge and requires proper techniques for management of data too.

4. Current status of research: With a Java program, permutation and combination of all the *Angalakshanas* have been generated, without any constraints for a single beat and the resulting output is more than **2 millions of dp vectors**. These constraints are related to those movements which are either physically not possible to perform or aesthetically not good movements to look at. For example the leg movements have 4 types: walking movements, leaps, circles and static postures. All 4 movements cannot be done at the same time by the performer and hence this is a physical constraint. An example of aesthetic constraint could be performing dance by looking at one corner only or looking down; which is never used for *Nritta*. Even after considering few such constraints, the result has been reduced to few thousands and needs further more methods to reduce the output to a manageable state. The research idea is at its infancy yet.

5. Conclusion: The tool developed could help in choreography. It can generate various permutations and combinations of *Angalakshanas* in *Bharata Natyam* by which many mundane tasks of a choreographer will be reduced. The choreographer shall not have to think which movements to give for *Nritta*, every time it occurs in a music number. This tool can quickly show an entire list of moves available for a particular beat hence voiding repetition of the choreographer's favorite move again and again and thus will create unusual choreography. It can also be used as a tool in dance education by introducing students to a variety of new dance movements and allowing them therefore to explore new ways to expressively combine those movements.

6. References

- [1] Scott de Lahunta, "The Future of Choreographic practice :the Choreographic Language Agent (CLA)", 2009, February
- [2] Mario Nakazawa and Andrea Paezold-Ruehi, "Dancing ,

Dance And Choreography; an Intelligent Nondeterministic Generator", ACM, 2009 , 30-34, April

- [3] Tom Calvert, "Animating Dance", ACM, 2007, May
- [4] Tom Calvert ,Lars Wilke , Rhonda Ryman and Ilene fox, "Applications of Computers to Dance", IEEE, 2005, 6-12, July
- [5] James H Bradford and Paulette cote- Laurence, "An Application of Artificial Intelligence to the choreography of Dance", Computers and the humanities, 1995, 29, 233-240
- [6] Soumyadeep Paul, Sudipta N. Sinha and Amitabha Mukherjee, "Virtual Kathakali: Gesture driven metamorphosis", Centre for Robotics, IIT Kanpur, http://www.iitk.ac.in/robotics/project_lists/gesture/kathakali/
- [7] Sumant Narayan Pattanaik, "A Stylised Model for Animating Bharata Natyam, an Indian Classical Dance Form" , *Computers in art, design, and animation* , John ansdown, Rae A. Earnshaw (eds.) Springer-verlag New York Inc., USA, 1989 .
- [8] Labanotation, The University Of Frankfurt, <http://www.gotan.ch/labnotation.html>
- [9] <http://user.unifrankfurt.de/~griesbec/LABANE.HTML>
- [10] <http://user.unifrankfurt.de/~griesbec/CHOREOE.HTML#LabanPad>
- [11] <http://user.unifrankfurt.de/~griesbec/LIMEE.HTML>
- [12] <http://user.unifrankfurt.de/~griesbec/LABPDE.HTML>
- [13] <http://web.bham.ac.uk/calaban/frame.htm>
- [14] <http://www.lifeforms.com/products/index.htm>
- [15] Labanwriter, The Ohio State University, http://dance.osu.edu/3_research_gallery/Labanotation_history.html
- [16] Annemette P Karpen's website, http://www.klassisk-indiskdans.dk/13bharata%20natyam_kathakali_odissi_kathak_dans_koebenhavn_klassisk%20indisk%20dans_kas_trupvej%2098%20a_indisk%20musikforening_chau.htm

Haploid vs Diploid Genome in Genetic Algorithms for TSP

Rakesh Kumar¹, Jyotishree²

1 Associate Professor, Department of Computer Science & Application, Kurukshetra University, Kurukshetra
Email: rsagwal@rediffmail.com

2 Assistant Professor, Department of Computer Science & Application, Guru Nanak Girls College, Yamuna Nagar
Email: jyotishreer@gmail.com

Abstract: There exist all types of organisms in nature – haploid, diploid and multiploid. Maximum research works in Genetic Algorithms are carried out using haploids. Diploidy and dominance have not been given due weightage although in maximum complex systems nature uses them. The paper illustrates the previous research work in diploidy and dominance. In this paper, a Genetic Algorithm is proposed to solve Traveling Salesman Problem (TSP) using haploid and diploid genome and to compare their performance in terms of cost and time.

Keywords— Genetic algorithms, Diploidy, Dominance

I. INTRODUCTION

Genetic Algorithms are considered to be apt for problem solving involving search. In contrast to other conventional search alternatives, they can be applied to most problems, just focusing on good function specification and a good choice of representation and interpretation. Moreover, the exponentially increasing speed/cost ratio of computers makes them a choice to consider for any search problem. They are based on Darwin's principle of 'Survival of fittest'. Most of the research works in genetic algorithms make use of haploid genomes, which contain one allele at each locus. But in nature, many biological organisms, including humans, have diploid genomes having two alleles at each locus and even some organisms have multiploid genomes having two or more alleles at each locus. This paper reviews various implementations of diploidy in different applications and implements diploidy in genetic algorithms to solve the traveling salesman problem.

II. DIPLOID GENOME AND DOMINANCE

In natural systems, the total genetic package is called genotype and organism formed by interaction of total genetic package with its environment is called phenotype. Each individual's genotype consists of a set of chromosomes having genes which may take some value called allele [10]. Each gene corresponds to a parameter of optimization problem. The simplest genotype in nature is haploid which has single chromosome. Diploid genome has

two chromosomes that consist of two set of alleles representing different phenotypic properties.

Genotype of diploid organisms contains double the amount of information for same function than the haploids. This leads to lots of redundant information which is eliminated by the use of genetic operator – Dominance. At a locus, one allele takes precedence over other alleles. Dominant alleles are expressed and denoted by capital letters and recessive ones by small letters in the phenotype. Dominance can be referred to as genotype to phenotype mapping or genotype reduction mapping [7]. It could be represented as:

$$\begin{array}{l} \text{AbCDe} \\ \text{aBCde} \end{array} \rightarrow \text{ABCDe}$$

Diploidy and Dominance clearly state that double information in genotype is reduced by half in its phenotypic representation. Existence of redundant information in chromosomes and then its elimination leads to a thought provoking question. Why does nature keep double information in genotype and utilizes half of the information in phenotype? At first, this redundancy of information seems to be wasteful. But, it is hard fact that nature is not spendthrift. There must be some good reason behind the existence of diploidy and dominance in nature and keeping redundant information in genotype.

Diploidy provides a mechanism for remembering alleles and allele combinations that were previously useful and that dominance provides an operator to shield those remembered alleles from harmful selection in current hostile environment [7]. This genetic memory of diploidy stores the information regarding multiple solutions, but only one dominant solution is expressed in phenotype. Redundant information is carried along to next generation. Dominance or non-dominance of a particular allele is itself under genetic control and evolves.

Diploidy increases diversity in GAs by allowing recessive genes to survive in a population and become active at some later time when changes in the environment make them more desirable. One drawback of diploidy is that the mechanics of a diploid GA requires twice as much computational effort as the mechanics of a haploid GA because we have twice as many alleles to deal with [16].

III. HISTORICAL BACKGROUND

In 1967, Bagley used concept of diploidy to model a population of individuals with dominance map, thereby carrying hidden trait without expressing it [1]. Bagley added an evolvable dominance value to each gene. In 1967, Rosenberg simulated the evolution of a simple biochemical system in which single-celled organisms capable of producing enzymes were represented in diploid fashion and were evolved over time to produce appropriate chemical concentrations. Any dominance effect was result of presence or absence of particular enzyme [22].

In 1971, Hollstein used a dominance schedule and suggested that diploidy did not offer a significant advantage to fitness. He described two simple evolving dominance mechanisms [9]. In the first scheme, each binary gene was described by two genes, a modifier gene and functional gene. Hollstein further replaced this two-locus evolving dominance by simpler one-locus scheme by introducing third allele at each locus and named it as triallelic scheme. This triallelic scheme was analysed for its steady state performance by Holland in 1975 and it turned out to be clearest, simplest Hollstein-Holland triallelic scheme for artificial genetic search. It combined dominance map and allele information at a single position [10].

Many experiments on function optimization were carried out by A. Brindle in 1981 with different dominance schemes. She did not consider artificial dominance and diploidy as taken in earlier experiments and developed six new dominance schemes [2]. In 1987, D.E. Goldberg and R.E. Smith used diploid representations and a dominance operator in GA's to improve performance of non-stationary problems in function optimization. They used three schemes: a simple haploid GA, a diploid GA with a fixed dominance map (1 dominates 0) and applied them to a 17-object, blind, nonstationary 0-1 knapsack problem where the weight constraint is varied in time as a periodic step function [6]. They proved the superiority of diploidy over haploidy in a nonstationary knapsack problem.

In 1992, R.E. Smith & D.E. Goldberg extended their research and showed that a diploid GA maintained extra diversity at loci where alternative alleles were emphasized in the recent past [18]. In 1994, F. Greene used diploid/dominance in genetic search. Diploid chromosomes were computed separately and were evaluated to produce two intermediate phenotypes. Mapping function was called dominance map or dominance function and fitness was

referred to as sub-fitness [8]. Experiment was performed using diploid chromosome C++ object compatible to Genitor and arithmetic crossover was implemented on it and they identified the changing global optima.

In 1994, C. Ryan avoided the use of dominance altogether and introduced two new schemes - additive diploidy and polygenic inheritance. Additive diploidy scheme excelled in non-binary GAs as it can be applied to GAs with any number of phenotypes. The implementation of high level diploidy is referred to as the degree of Nness, where N is the number of the last phenotype [15]. In 1995, K.P. Ng and K.C. Wong proposed dominance change mechanism in which dominance relationships between alleles could change over time. They extended the multiallele approach to dominance computation by adding a fourth value for a recessive 0. Thus 1 dominates 0 and 0 and 0 dominates i and o. When both allele values for a gene are dominant or recessive, then one of the two values is chosen randomly to be the dominant value. They also suggested that the dominance of all of the components in the genome should be reversed when the fitness value of an individual falls by 20% or more between generations and system is not suitable for domains where changes are small [13].

In 1996, Calabretta et al compared the behavior of haploid and diploid populations of ecological neural networks in fixed and changing environments. They showed that diploid genotypes were better than haploid ones in terms of fitness and diploid genotypes retained better changes in environment. They analysed the effect of mutation on both type of populations [3]. They concluded that diploids had lower average fitness but higher peak fitness than haploids. In 1996, E. Collingwood, D. Corne & P. Ross studied the use of multiploidy in GAs for two known test problems namely, the Indecisive(k) problem and the max problem. In their multiploid model, they used p chromosomes and a simple mask which specified dominant gene at a locus in each chromosome and further this mask helped to derive the phenotype. On testing the two problems at same population size, they analyzed that the multiploid algorithms outperformed haploid algorithms [5]. Multiploid GA was able to recover from early genetic drift, thereby good genes managed to remain in population, shielded from harmful over-selection of bad genes.

In 1997, Calabretta et al used a 2-bit dominance modifier gene for each locus apart from structural genes expressing neural network [4]. They compared the adaptation ability of haploid and diploid individuals in varying environment and found that diploid populations performed better and were able to tolerate sudden environment changes, thus exhibiting less reduction in fitness. In 1998, J. Lewis, E. Hart and G. Ritchie tested various diploid algorithms, with and without mechanisms for dominance change on two variations of nonstationary problems. Comparison showed that diploid scheme did not perform well and on adding

dominance change mechanism, performance improved significantly. Further, extending the additive dominance scheme with change improved the performance considerably [12]. They concluded that some form of dominance mechanism is needed along with diploidy to allow flexible response to change.

In 2002, S. Ayse, Yilmaz, Annie S. Wu, proposed a new diploid scheme without dominance on integer representation. In their research, they evolved all diploid individuals without using any haploid stage and compared performance for TSP as their test problem. They concluded that simple haploid GA outperformed diploid GA [21]. In 2003, C. Ryan, J.J. Collins, D. Wallin extended their work and proposed the Shades scheme – a new version of haploidy that incorporates the characteristics of diploid scheme in haploid genetic algorithms but with lesser cost. Performance of Shades scheme was analyzed and compared to two diploid schemes – tri-allelic and Dominance change mechanism scheme in two dynamic problems domains. Shades-3 outperformed both the diploid schemes in both Osmera's dynamic problem and constrained knapsack problem. Shades-2 outperformed shades -3 in knapsack problem [16].

In 2003, Robert Schafer presented a GA protocol as a tool to approach dynamic systems having reciprocal individual-environment interaction and then applied on a model problem in which a population of simulated creatures lived and metabolized in a three-gas atmosphere [17]. In 2005, Shane Lee and Hefin Rowlands described a diploid genetic algorithm, which favoured robust local optima rather than a less robust global optimum in a problem space. Diploid chromosomes were created with two binary haploid chromosomes, which were then used to create a schema. The schema was then used to measure the fitness of a family of solutions. [11].

In 2007, Shengxiang Yang proposed an adaptive dominance learning scheme for diploid genetic algorithms in dynamic environments. In this scheme, the genotype to phenotype mapping in each gene locus was controlled by a dominance probability [20]. The proposed dominance scheme was experimentally compared to two other schemes for diploid genetic algorithms and results validated the efficiency of the dominance learning scheme. Out of the two schemes, additive diploidy scheme proved to be better than the Ng-Wong dominance scheme. In 2009, Dan Simon utilized diploidy and dominance in genetic algorithms to improve performance in time-varying optimization problems. He used the scaled One Max problem to provide additional theoretical basis for the superior time-varying performance of diploid GAs. The analysis confirmed that diploidy increases diversity, and provided some quantitative results for diversity increase as a function of the GA population characteristics [19].

IV. ALGORITHM

The Travelling Salesman Problem (TSP) is a classical combinatorial optimization problem, which is known to be NP-Hard problem. The problem is to find the shortest tour or Hamiltonian path through a set of N vertices so that each vertex is visited exactly once [14]. To find an optimal solution involves searching in a solution space that grows exponentially with number of cities. So, certain kind of heuristic techniques are applied to reduce the search space and direct search to the areas with the highest probability of good solutions. One such heuristics technique is genetic algorithms (GAs).

The problem is solved under following assumptions:

- Each city is connected to every other city,
- Each city has to be visited exactly once,
- The salesman's tour starts and ends at the same city.

Based on the above assumptions, a simple genetic algorithm is formulated to solve the problem.

GA-for-tsp(N,M,GP)

[N is number of cities and M is number of maximum generations, GP is generation pool]

```
1 Begin
2   0 ← i
3   Create an initial population P(i) of GP
   chromosomes having length N.
4   Evaluate the fitness of each chromosome in P(i).
5   While i < M do
6       Perform selection i.e. choose at random a
       pair of parents from P(i).
7       Exchange strings by crossover to create
       two offsprings.
8       Insert offsprings in P(i+1)
9       Repeat steps 6 to 8 until P(i+1) is full
10      Replace P(i) with P(i+1).
11      Evaluate the fitness of each chromosome
       in P(i+1)
12  end
13  Final result is best chromosome created during
   the search.
14 End
```

V. SIMULATION AND ANALYSIS

The algorithm is further coded in MATLAB for its implementation using both haploid and diploid genome set. The code was implemented first for 10 cities. The cost of different paths was computed for fifteen consecutive runs and then compared.

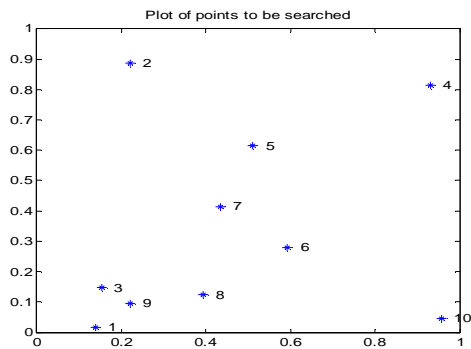


Figure 1

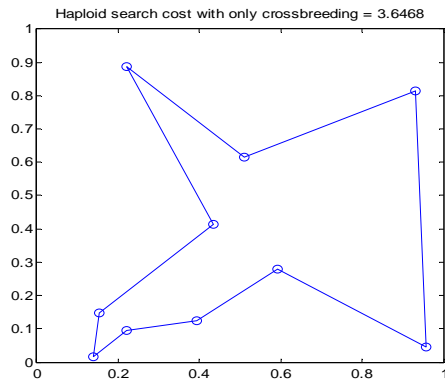


Figure 2

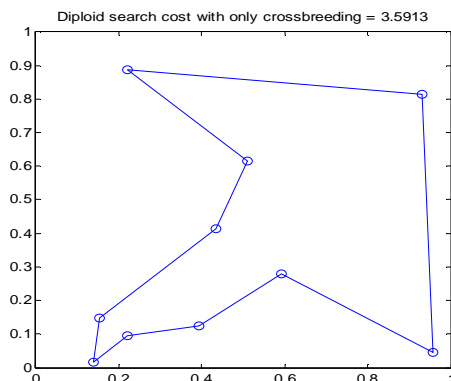


Figure 3

The implementation was then carried out for 50 cities and the results were compared. It was observed that in majority of runs both in case of 10 cities and 50 cities, diploid genome resulted in better results than haploid genome. The cost of path of final result using diploid genome was found to be less than that computed with haploid genome. Moreover, computational time was also found to be less in case of diploid chromosomes. Comparison of cost and time for different cases is illustrated in following figures.

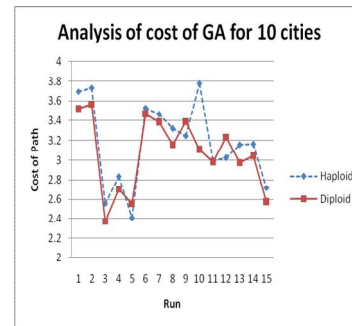


Figure 4

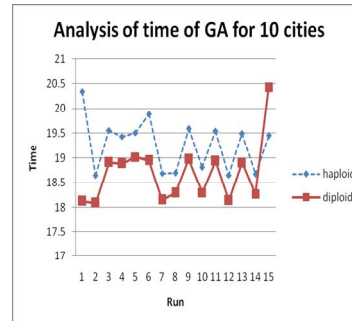


Figure 5

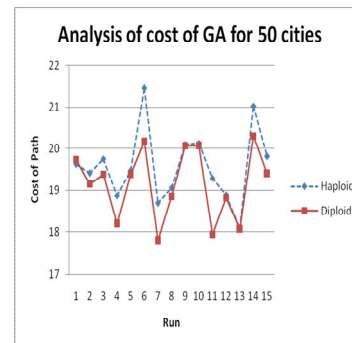


Figure 6

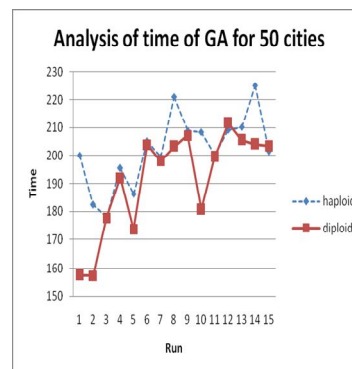


Figure 7

VI. CONCLUSION

By comparing haploid and diploid implementation of genetic algorithm, it has been shown that the genetic algorithm with diploid chromosomes performs better than the genetic algorithm with haploid chromosome. The experimental results show that the diploid GA can achieve faster response and is easy to implement. In continuation with the research work, it is proposed to develop a genetic algorithm using crossover probabilities and different crossover points to evaluate the performance in each case.

REFERENCES

- [1] Bagley, J.D., The behavior of adaptive systems which employ genetic and correlation algorithms. (Doctoral dissertation, University of Michigan) Dissertation Abstracts International 28(12), 5106B (University Microfilms No. 68-7556), 1967.
- [2] Brindle, A., Genetic algorithms for function optimization. Unpublished doctoral dissertation, University of Alberta, Edmonton, 1981.
- [3] Calabretta, R., Galbiati, R., Nolfi, S. & Parisi, D., Two is Better than One: a Diploid Genotype for Neural Networks. *Neural Processing Letters*, 4, 1-7, 1996.
- [4] Calabretta, R., Galbiati, R., Nolfi, S. & Parisi, D., "Investigating the role of diploidy in simulated populations of evolving individuals." *Electronic Proceedings of the 1997 European Conference on Artificial Life*, 1997.
- [5] Collingwood, E., Come, D. & Ross, P., "Useful Diversity via Multiploidy," *Proceedings Of the IEEE International Conference on Evolutionary Computing*, Nagoya, Japan, pp 810-813, 1996.
- [6] D. Goldberg and R. Smith, Nonstationary function optimization using genetic algorithms with dominance and diploidy, *Proceedings of Second International Conference on Genetic Algorithms*, Cambridge, MA, 59-68, 1987.
- [7] Goldberg, D. E., *Genetic algorithms in search, optimisation, and machine learning*. Addison Wesley Longman, Inc., ISBN 0-201-15767-5, 1989.
- [8] F. Greene, A method for utilizing diploid and dominance in genetic search, *Proceedings of the First IEEE Conference on Evolutionary Computation*, IEEE Press, pp. 439-444, 1994.
- [9] R. Hollstien, Artificial genetic adaptation in computer control systems, doctoral dissertation, University of Michigan, Ann Arbor, MI, Dissertation Abstracts International, 32(3) 1510B (University Microfilms No. 71-23,773), 1971.
- [10] Holland, J., *Adaptation in natural and artificial systems*, University of Michigan Press, Ann Arbor, 1975.
- [11] Shane Lee And Hefin Rowlands, Finding Robust Optima With A Diploid Genetic Algorithm, *I.J. Of Simulation* Vol. 6 No 9 73 Issn 1473-804x Online, 1473-8031 Print, 2005
- [12] J. Lewis, E. Hart and G. Ritchie. A comparison of dominance mechanisms and simple mutation on non-stationary problems. *PPSN V*, pp. 139-148, 1998.
- [13] K. P. Ng and K. C. Wong. A new diploid scheme and dominance change mechanism for non-stationary function optimisation. *Proceedings Of the 6th Int. Conf. on Genetic Algorithms*, pp. 159-166, 1995.
- [14] M. Perling, GeneTS: A Relational-Functional Genetic Algorithm for the Traveling Salesman Problem, Technical Report, Universitat Kaiserslautern, ISSN 09460071, August 1997
- [15] C. Ryan. The degree of oneness. *Proceedings of the 1994 ECAI Workshop on Genetic Algorithms*, Springer-Verlag, 1994.
- [16] C. Ryan, J.J. Collins, D. Wallin, Non-stationary function optimization using polygenic inheritance, in: *Lecture Notes in Computer Science*, vol. 2724, pp. 1320-1331, 2003.
- [17] Robert Schafer, Using a diploid genetic algorithm to create and maintain a complex system in dynamic Equilibrium, *Genetic Algorithms and Genetic Programming at Stanford 2003*, Book of Student Papers from John Koza's Course at Stanford on Genetic Algorithms and Genetic Programming, pp. 179-186, 2003.
- [18] Smith, R. E. & Goldberg, D. E. (1992). Diploidy and Dominance in Artificial Genetic Search. *Complex Systems*, 6(3), 251-285, 1992.
- [19] Dan Simon, An Analysis of Diploidy and Dominance in Genetic Algorithms, *International Conference on Computer, Communication, Control and Information Technology*, West Bengal, India, 2009
- [20] Shengxiang Yang, Learning the dominance in diploid genetic algorithms for changing optimization problems. *Proceedings of the 2nd International Symposium on Intelligence Computation and Applications*, pp. 157-162, 2007.
- [21] Ayse S. Yilmaz, Annie S. Wu, A comparison of Haploidy and Diploidy without Dominance on Integer representations, *Proceedings of the 17th International Symposium on Computer and Information Sciences*, Orlando, FL, pp. 242-248, 2002.
- [22] Rosenberg, R.S., "Simulation of Genetic Populations with Biochemical Properties", PhD Thesis, University of Michigan, 1967.

Context Based Personalized Search Engine For Online Learning

Dr. Ritu Soni , Prof. & Head, DCSA, ritusoniynr@yahoo.com

Mrs. Preeti Bakshi, Lect. Compuret Science, preetisegan03@gmail.com

GNG College, Santpura, Haryana, INDIA-13

Abstract: The WWW serves as a huge widely distributed, global information service centre for acquiring on line knowledge. But there are lots of challenges in finding up the effective, relevant information. During the past few years, it has become apparent that there is an urgent need for tools that guide Web users in their information finding and navigational activities. One of the failings of traditional information retrieval models is attributed to the isolation of queries from the context in which they occur. Mounting a system that can understand the user's seeking needs beforehand and then presents a list of ranked links to the web pages most probably having the relevant information according to user's context to lift up information retrieval process that meets the user's requirement in a effective manner. With this objective, Context Based Personalized Search Engine for online Learning supports the hypothesis that the relevance of search results can be improved through the inclusion of contextual information. This contextual information can be obtained by an automated analysis of the local information space surrounding a given candidate page with the local space defined by the navigational structure inherent the Web.

Keywords: *WWW, web personalization, Context Management, Conceptual Architecture, CFL, IR*

1. Introduction: In order to satisfy human quest for knowledge, self paced, On-line learning is gaining its importance. The World Wide Web [1] serves as a huge widely distributed, global information service centre for news, advertisements, customer information, financial management, education, government, e- commerce and many other information services. The Web also contains a rich and dynamic collection of hyperlink information and Web page access and usage information. . Given that there is this vast and ever growing amount of information, how does the average user quickly find what she is looking for a task in which the present day search engines don't seem to help much! This is because of high complexity of Web page collection, lack of unifying structure with respect to indexing by category or by title, author, cover page, table of contents etc, highly dynamic nature, broad diversity of user communities that uses only 1% of relevant information, non-effective & inefficient simple keyword based search engine for Web resource discovery and the requirement for adding value to e-services on the Web, necessity towards the creation of loyal visitors customers for a Web site also adds to the disadvantage of Web usage.[2][3]

However, based on the above observations, the web also poses great challenges for effective resource and knowledge discovery. How can the portion of the Web that is truly relevant to the user's interests be determined? How can we find high quality Web pages on a specific topic? These

challenges have promoted research into effective discovery and efficient use of resources on the internet. One possible approach is to personalize [4] the web space. "Web personalization is simply defined as a task of making Web based information system adaptive to the needs and interest of users". In other words a system is created which responds to user queries by potentially aggregating information from several sources in a manner which is dependent on who the user is.[5]

The techniques of mining & the user's access patterns have been applied in a wide range of application including Web search engine and Web personalization.

Search engines [6] are generally affected by problems ambiguity and results ordered by website popularity rather than user interests. As most of the present search engines treat the search requests in isolation, they create an index of words within documents and return a rank list of documents in respond to user queries. However only few of the results, returned by the search engines may be valuable. The document is valuable to the user or not depends upon the context of the query asked.

It is very important to have dedicated search engine for specific study. Such an application will save user's time from searching from unstructured web.

"The proposed model generates links to the user's query by understanding the context in which query is initiated and then personalizing user's request according to the user's interest area so that the information needs of the user is satisfied" is the idea behind the Proposed Model.

The idea of Personalizing the Information Retrieval in User context is achieved by considering and combining the two different but related concepts:

- a) Personalization and
- b) Context Management.[7]

Personalization is a Process of presenting the right information to the right user at the right moment. [8] It is achieved through creating user profile, which gives the description of user interest, preference etc. Information about the user can be collected in two ways: i) Explicitly, for example asking for the feedback such preferences or ratings; and ii) Implicitly, for example observing user behavior such as the time spent reading online document. The presented model follows the first way i.e. it creates the user profiles by explicitly gathering the information about the user preferences. Based upon this personalization information search for requested query is contextualized and system presents better results to the user. It is common knowledge that several forms of context exist in the area.

2. Objective of Study

During the past few years, it has become apparent that there is an urgent need for tools that guide Web users in their information finding and navigational activities. One of the failings of traditional information retrieval models is attributed to the isolation of queries from the context in which they occur.

Developing a system that can understand the user's seeking needs beforehand and then presents a list of ranked links to the web pages most probably having the relevant information according to user's context to lift up information retrieval process that meets the user's requirement in an effective manner.

The above stated is the main objective of our work. Our work supports the hypothesis that the relevance of search results can be improved through the inclusion of contextual information. This contextual information can be obtained by an automated analysis of the local information space surrounding a given candidate page with the local space defined by the navigational structure inherent in the Web.

To obtain the above stated goal our work exploits the contextual information and smoothly integrates it into the personalization of information retrieval. The idea of contextual personalization [9] proposed and developed here responds to the fact that human preferences are complex, multiple, heterogeneous, changing even contradictory and should be understood in context with the user's goal and task

at hand. Context is a difficult notion to grasp and capture in a software system. Several researchers have tried to categorize context-aware applications and features including the Contextual sensing, Contextual adaptation, Contextual resource discovery and Contextual augmentation. [10]

We have focused our efforts on the major topic of information search and retrieval system by restricting it in context of user's usage access patterns while browsing the Web.

The approach behind this work is alleviating information overload offered by the Web on the user and making the Web a friendlier environment for its individual user and hence creating trustworthy relationship between the Web-site and the visitor-customer.

3. Proposed Model

The model presented here is concerned with exploiting semantic, ontology based contextual information, specifically aimed towards its use in Personalization for information retrieval. The goal of the presented model is to endow personalized system with the capability to filter and focus their knowledge about user preferences on the semantic context of ongoing user activities.

This paper aims at creating a system, that is (semi) automatically tailored, for the content delivered to the user, from a web site. The proposed system constitutes both the contents, as well as the users' interaction. By applying mining methodologies on the Web data one can discover or understand the user's interest area or his/her habits of browsing the web.

The study involves the process of getting user's interest and needs, analyzing and organizing this valuable information in the context in accordance to user's requirement, and then providing the user, the result in the form of links that connects to the Web pages containing the needed detail.

3.1 Components for proposed system

Conceptual Architecture of the proposed system has components for front-end tools for semantic based web services & administration tools; required functionality tools and data storages & ontology with proper security.

The core components for proposed system provide the required functionality for semantic web enabled web services [11] for on line learning. These components are as follows:

Administrator: It controls, co-ordinates & administer all other components.

Message Transformation: This component supports the semantic transformations at different levels of data, logics & rules, protocol engines etc.

Fig 1 Conceptual Architecture

Web Search management: This component enables the search of a particular web page in a particular user profiles in administration process.

Auditing / Tracking: It provides trust to the different users through administrator.

User Manager: It has complete user profile.

Security: This component provides security for authorization & authentication of user.

Web Services & Goal Manager: This component allows managers to standardize & simplify data access while hiding the complexity of processes.

Protocol Engine: This allows message passing over network.

Deployment: It makes Web services available for the user in different platforms.

Adapter: They provide the adaptability to the future technologies so as to reuse the objects.

Transport: It provides network and protocol connectivity.

Fig 2 CFL Semantic web architect

The proposed model can also be explained through different layers of CFL Semantic web architect as described above.

3.2 Model Design:

The proposed model is structured as a search engine that makes the use of consideration described above. It records the user preferences in the form of links navigated while browsing the site in the logs and create user profiles for their regular users. Whenever that particular visits the site again these logs are accessed to personalize the search in user's context.

The proposed model is structured in two views one is User View and second is Administrator View. The user has an access only to User View. It provides trust on the web site with the help of authorization and authentication rules. Users can retrieve information according to their requirements, from the data stored. This is possible by user manager component. The Administrator View is only accessible by the administrator for the management of the model designed. Functional and technical tools are developed using asp.net.

3.4 Model Requirements:

- The proposed model is design in two views ie Users view and Administrator of the Site. The user provides an access only to the User's View. It is

only the Administrator who can make the changes in the Site functioning and in Data Base.

- The Proposed Model is based on the explicit approach of personalizing search in user Context. Here user is required to mark and submit his/her interests in the form of preferred links; it is not done automatically by the system. Thus the process of extracting user interest is user dependent.

- Only the authorized user can have an access to the system. However any kind of user can register him/her as the authorized user just by filling up a registration form. There will be no charges or fee for the membership.

Boundary Conditions:

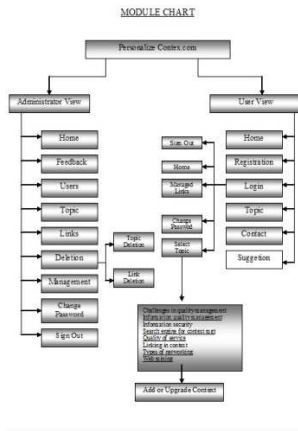
The proposed model is design in two views. One is for the Users and Second is for the Administrator of the Site. The user provides an access only to the User's View. It is only the Administrator who can make the changes in the Site functioning and in Data Base.

- A number of techniques has been evolved and developed in the area of Context Management as explained in the chapter 4. The Proposed Model is based on the Explicit approach of Personalizing search in user Context i.e. it requires user to enter his/her interest to the system. In the Proposed model user is required to mark and submit his/her interests in the form of preferred links, it is not done automatically by the system. Thus the process of extracting user interest is user dependent.

- Only the authorized can have an access to the system. It requires that users must have an account with the system to avail its services. However any kind of user can register him/her as the authorized user just by filling up a registration form. There will be no charges or fee for the membership.

The hypothesis depicted in the given study is applied only to the limited volume of the data, to show the effect that how context management enhance the results of information retrieval process and provides benefits to the user.

Module chart of the Proposed System

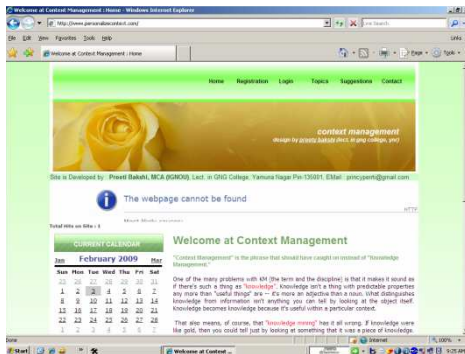


4. Implementation:

Taking the consideration of personalizing search in user's context as explained in the previous section of the study the proposed model is implemented as a Web search engine.

The Implemented model has two views. One is **User view** and the Second is **Administrator View**. The View available for the visitors is the User View. The Administrator View having different domain name, is kept hidden from the users and is accessible to administrator only. The administrator can use this view for different analysis purposes or to upgrade the system services or can use it for different decision making purposes.

4.1 The User View: The processes under user view are explained below. This is the start page of site.



Brief functioning of user view: The option **Topics** provides only the general view of the topics and underlying links. It is only the general view, search will not initiate from here. The system will open only to the registered user through the login process. After this search proceeds from the User's Welcome Window. From here user can select the required topic and can get the required information

by navigating the links. The user is supposed to manage its profile by **Add/Update account**. This will help in personalizing the search for the next session. Some more options are available to view the managed links and to contact to the Administrator. User can close his account by sign out. Each screen has the link to go to the home page. User can also place its search experience through the suggestion link.

The user can navigate the site through the following links:

Home: To come to the home page of the Site.

Registration: To register yourself as a authenticated user of the Site.

Login: To get the access of the site service.

Topics: To see the topics and the related links about which the system has store the data.

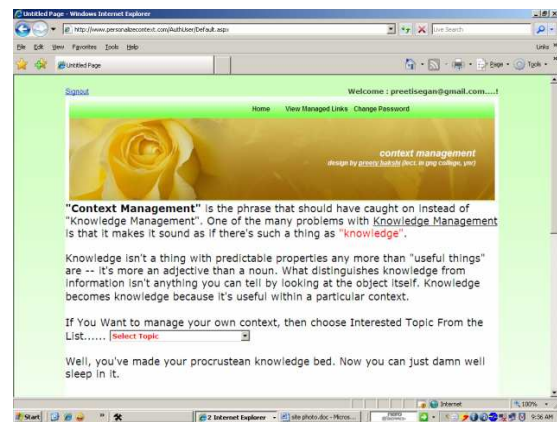
Suggestions: To get the feedback from the user.

Contact: Help to contact with the administrator.

Details of some of the user view's components are as follows:

• User's Welcome Window:

This window appears when the existing user performs the login process. Actually it is the home page of the user's window.



Process Description:

After the user has successfully finished the Login process, the system will show the User's welcome window to the user. From here the users can select the topic of their interests and can find the information for which they are looking for. This window also provides some more function to the user to manage his/her account or interests. The functions are listed below:

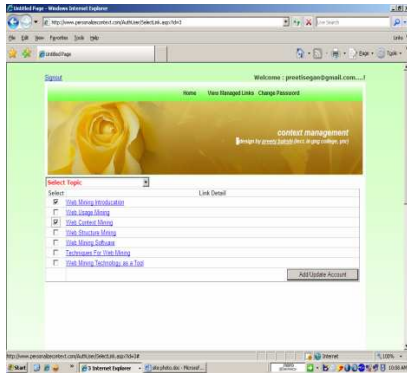
Home: It will show the same window presented above.

View Managed Links: This will show the list of topics and links navigated by the users in the previous session.

Change Password: This will lets user to change the password.

- **Link Navigation:**

The window shows the list of links related to a particular topic.



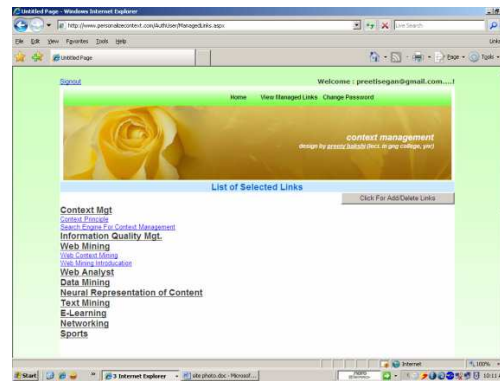
Process Description:

From here the user can select any number of links and navigate them to find the information he/she seeking for. To manage the links according to the user's context the system requires that the particular user mark the interested links and then add or update this information to his/her profile by clicking on the **Add/Update Account** button on the lower left corner of the window. The system creates or maintains the user profile for the purpose of personalizing the search in the user's context to gain the benefit in the Information Retrieval (IR) Process. When the same user visits the same topic area again, then these profiles help the system in providing the user a list of rank links according to his/her interests.

- **View Managed Links:**

This window presents list of topics and links marked by the user as interests in the previous session of the visit to the site.

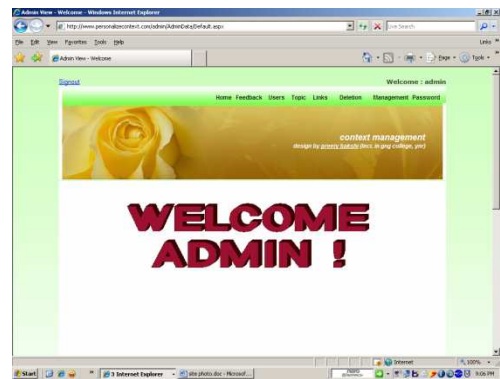
This process shows a list of rank links marked by the user as interested links in previous session of his/her visit. The user can start his navigation process here. This provides an advantage to the user, as it saves lots of time of user spending in finding the required or interested links he/she wants to check out.



Process description:

4.2 The Administrator View: This view is accessed by the administrator only. He can make changes to the system to meet the users' requirement. The processes under administrator view are listed below.

This form is the administrator start page. This page will display after the successful Login process. Administrator View also presents various functions to the Administrator that is listed below.



Home: This will show the Administrator's home page.

Feedback: This will show the suggestions given by the users.

Users: This will show the list of the users of the site.

Topic: This will add a topic to the topic list.

Links: This will add link to the specified topic.

Deletion: This will delete a topic or a link under a topic.

Management: This will give the list of user and the links navigated and marked by them as their interest areas.

Password: This will change the password of the Administrator.

5. Conclusion

Context is an increasingly common notion in Information Retrieval. Several authors in the IR field have explored approaches that are similar to one presented. The distinctive aspect in our

approach is the use of semantic concepts, rather than plain terms (i.e. keywords), for the representation of these contextual meanings, and the exploitation of explicit ontology-based information attached to the concepts, available in a knowledge base. This extra, formal information allows determining the set of concepts that can be properly attributed to the context, in a more accurate and reliable way (by analyzing explicit semantic relations).

This paper proposes an approach to improving the precision of search engines, whilst maintaining the level of recall achievable, by utilizing the local information space (as defined by linked pages) in providing a search context. To demonstrate the validity of the approach we have developed a simple model of search engine. This model- even though based on a very simple model of context - is sufficient to provide statistically significant enhancement of the relevance of search results. We are currently carrying out further tests to quantify more precisely the level of improvement that is achievable.

6. Future Scope

To improve this approach further, there are a number of directions for further work. The most significant of these is to develop an improved model of the relationship between information context and hyper-textually linked pages in a web environment. This model forms the foundations of the approach being adopted and is required for developing a more rigorously supported set of searching algorithms.

7. References

[1] World Wide Web Consortium (W3C) Semantic Web activity homepage.
<http://w3c.org/sw.rg/wc/robots.html>.
[2] Berners-Lee T. The Semantic Web and Challenges. W3C Web site Slideshow 2003.
URL: <http://www.w3.org/2003/Talks/01-sweb-tbl/Overview.html>
[3] Berners-Lee T. Standards, Semantics and Survival. *SIIA Upgrade* 2003:pp 6–10.
[4]. Sergio Flesca, Sergio Greco, Ester Zmpino, "Mining User Preferences, Page Count and Usage to Personalize WebSite Navigation", August 2005
[5] P.H Mylonasi, D. Vallet, P. Castell, M. Ferna and Y. Avrit Hi Si, "Personalized information retrieval based on context", 2008, The Knowledge engineering Review, Vol. 23:1,73-100, Cambridge university Press.
[6] Antoniou G. and von Harmelen F. *A Semantic Web Primer*. The MIT Press, Cambridge, Massachusetts, London, England 2004. ISBN 0-262-01210-3.

[7] Ajay D Kshemkalyani, George Samaras and Andrew Citron, "Context management and its applications to distributed transactions", December 1996
[8] Michael Sintek and Stefan Decker. Triple - a query, inference, and transformation language for the semantic Web. In Proceedings of the First International Semantic Web Conference on The Semantic Web, Sardinia, Italy, 9-12 June 2002. Springer-Verlag.
[9] Ian Horrocks and Sergio Tessaris. "Querying the semantic Web: A formal approach", In Proceedings of the 1st International Semantic Web Conference, Sardinia, Italy, 9-12 June 2002. Springer-Verlag.
[10] A Stephen J.H. Yang a, Jia Zhang b, Irene Y.L. Chen c, " JESS-enabled context elicitation system for providing context-aware Web services", The Journal of system and software 2007.
[12] D Fensel, C. Bussler, and A. Maedche, " Semantic web enabled services" , proceeding of International Semantic Web Conference 2002, LNCS, Springer, pages 1-2,2002
[13] A Ankolenkar, M. Burstein, T. C. Son, J. Hobbs, O.Lassila, D Martin, D McDermott, S. McIlraith, al et. "DAML-S : Semantic Markup Web Services" Online available at <http://www.daml.org/services/2001>
[14] M.Shaw and D Garlan: "Software Architectures" Prentice-Hall, 1996.
[15] G. Wiederhold. Mediators, "Architecture of future information systems". *IEEE Computer*, 25(3), 1992.
[16] D. Fensel and R. Groenboom. A software architecture for knowledge-based systems. "The Knowledge Engineering Review (KER)", 14(3), 1999.

SELF-HEALING IN WIRELESS ROUTING USING BACKBONE NODES

Urvi Sagar¹, Ashwani Kush²

²CSE, NIT KKR, email

¹Comp Sci Dept, University College, Kurukshetra University India, akush20@gmail.com

Abstract:

Wireless networking is a new emerging era. It has potential applications in extremely unpredictable and dynamic environments. Individuals and industries choose wireless because it allows flexibility of location, whether that means mobility, portability, or just ease of installation at a fixed point. A flat mobile ad hoc network has an inherent scalability limitation in terms of achievable network capacity. It is seen that when the network size increases, per node throughput of an ad hoc network rapidly decreases. This is due to the fact that in large scale networks, flat structure of networks results in long hop paths which are prone to breaks. The challenge of wireless communication is that, the environment that wireless communications travel through is unpredictable. Wireless networks that fix their own broken communication links may speed up their widespread acceptance. The changes made to the network architectures are resulting in new methods of application design for this medium. The long hop paths can be avoided by using backbone nodes concept. In this paper, a self healing scheme for large scale networks with mobile backbone nodes has been proposed.

Keywords: *MANET, routing, ADOV, Self healing network*

1.0 Introduction

There is tremendous technological advance in producing small and smart devices. The number of embedded devices in appliances and vehicles is increasing at a rapid rate. Thousands of such devices can be used for applications[1] like: environmental data collection, weather forecasting, measuring toxicity levels at hazardous sites etc. It is a natural consequence that such devices work in a collaborative way. However, users carry around many such smart devices and they are not fixed in the sense of a desktop computer. Hence, there is a need for networking such mobile devices without any infrastructural support. There is a growing demand of using networks of mobile devices[2] anywhere and anytime. Cellular Phones and Internet provide some solution, but Cellular phones work with infrastructural support like mobile phone towers and satellite communication. However, such support comes at a cost like pre-registration with a mobile service provider etc. In many situations, the Internet may not be an efficient solution. For example, a collection of people trying to communicate in a hotel or conference hall. Adhoc network provide a solution to these problems. An ad hoc network is a collection of autonomous nodes, which may move arbitrarily so that the topology changes frequently. In contrast to conventional wireless networks, the nodes in Mobile ad hoc network communicate using wireless links without any fixed network infrastructure and centralized administrative support. A node act both as source/destination for messages and as a switching or routing node. The purpose of an ad hoc network is to set up (possibly) a short-lived network for a

collection of nodes. If all the wireless nodes are within the transmission range of each other, routing is easy. Every node can listen to all transmissions. However, this is not true in most situations, due to short transmission range. Hence, most ad hoc networks are multi-hop [3]. A message from a source node must go through intermediate nodes to reach its destination. All nodes cooperate in delivering messages across the network. A major problem is ad hoc network is route stability as mobility has a significant effect on network integrity. Link failures lead to a considerable packet loss in data transmission. In this paper a new proposal based on backbone nodes has been introduced to make route stable and follow the concept of self healing. Rest of paper is organised as : Section 2 highlights major issues of ad hoc network, Section 3 gives a detailed survey of self healing networks with techniques, proposed scheme is part of section IV and results and discussion have been made in section V.

2.0 Major issues in Ad hoc networks [4,5]

- Most nodes in an ad hoc network are powered by batteries. Batteries cannot be recharged easily in many cases. Each node participates in two kinds of activities, sending and receiving messages useful for itself and forwarding messages for other nodes.
- Mobile communication is needed. Communication must take place in a terrain that makes wired communication difficult or impossible. A communication system must be deployed quickly.
- Communication facilities must be installed at low initial cost. The same information must be broadcast to many locations. Operates in a less controlled environment, so is more susceptible to interference, signal loss, noise, and eavesdropping.
- Network support for user mobility
- Efficient use of finite radio spectrum
- Integrated services (voice, data, multimedia)
- Maintaining quality of service over unreliable links
- Security
- Cost efficiency
- The issue of the reliability

3.0 Self Healing Network

In developing broadband digital networks, a short service-outage such as a link failure or a node failure can cause a serious impairment of network services. It is due to the volume of network traffic carried by a single link or node. Moreover, the outage can stimulate end users to try to re-establish their connections within a short time. The retrials, however, make the problem worse because the connection establishment increases the traffic volume further. Fast restoration from a network failure becomes a critical issue in deploying high-speed networks. Self healing algorithms have been recognized as a major mechanism for providing the fast restoration. A self-healing system [6] should recover from the abnormal state and return to the normal state, and should start functioning as it was prior to failure. One of the key issues associated with self-healing networks is to optimize the networks while expecting reasonable network failures [6,7,8].

Self-healing network (SHN) [9] is designed to support transmission of messages across multiple nodes while also protecting against recursive node and process failures. It will automatically recover itself after a failure occurs. The problem of self-healing is in networks that are reconfigurable in the sense that they can change their topology during an attack. One goal is to maintain connectivity in these networks, even in the presence of repeated adversarial node deletion. Modern computer systems are approaching scales of billions of components. Such systems are less akin to a traditional engineering enterprise such as a bridge, and more akin to a living organism in terms of complexity. A railway overbridge must be designed in such a way that, key components never fail, since there is no way for the bridge to automatically recover from system failure. In contrast, a living organism can not be designed so that no component ever fails: there are simply too many components. For example, skin can be cut and still heal. Designing skin that can heal is much more practical than designing skin that is completely rigid to attack. Unfortunately, current algorithms ensure robustness in computer networks through hardening individual components or, at best, adding lots of redundant components [10].

Critical issues [11] in self-healing systems typically include ; Maintenance of system health, recovery processes to return the state from an unhealthy state to a health one. Self-healing components or systems typically have the following characteristics [11] : (a) perform the productive operations of the system, (b) coordinate the activities of the different agents, (c) control and audit performance, (d) adapt to external and internal changes and (e) have policies to determine the overall purpose of the system. Most of the self-healing concepts are still in very early stages; still some possible areas explored are Grid computing, software agents, middleware computing, ad hoc networks. Emphasis here is on ad hoc network self healing characteristic. This section provides an analysis of various schemes that can be used as self healing schemes.

a) Self Healing in Routing

The most promising developments in the area of self-healing wireless networks are ad hoc networks. They are decentralized, self-organizing, and automatically reconfigure without human intervention in the event of degraded or broken communication links between transceivers. Automated network analysis through link and route discovery and evaluation are the distinguishing features of self-healing network algorithms. Through discovery, networks establish one or more routes between the originator and the recipient of a message. Through evaluation, networks detect route failures, trigger renewed discovery, and—in some cases—select the best route available for a message. Because discovery and route evaluation consume network capacity, careful use of both processes is important to achieving good network performance.

b) Self healing in RF

Environmental radio-frequency (RF)[12] “noise” produced by powerful motors, other wireless devices, microwaves—and even the moisture content in the air—can make wireless communication unreliable. Despite early problems in overcoming this pitfall, the newest developments in self-healing wireless networks are solving the problem by capitalizing on the inherent broadcast properties of RF transmission. The changes made to the network architectures are resulting in new methods of application design for this medium.

c) Self healing in Power efficiency

As the network is always on, conserving power is more difficult. One solution is **On-demand discovery**[11]. It establishes only the routes that are requested by higher-layer software. On-demand discovery networks are only “on” when called for. This allows nodes to conserve power and bandwidth and keeps the network fairly free of traffic. If, between transmissions, the link quality between nodes has degraded, however, on-demand networks can take longer to reconfigure and, thus, to deliver a message. Once routes have been established, they must generally be maintained in the presence of failing equipment, changing environmental conditions, interference, etc. This maintenance may also be proactive or on-demand. Another solution can be **Single-path routing**[11]. As for routing, network algorithms that choose single-path routing, as the name suggests, single out a specific route for a given source-destination pair. Sometimes, the entire end-to-end route is predetermined. Sometimes, only the next “hop” is known. The advantage of this type of routing is that it cuts down on traffic, bandwidth use, and power use. If only one node at a time needs to receive the packet, others can stop listening after they hear that they’re not the recipient.

3.1 Self-Healing Technologies

Dynamic Source Routing (DSR)

DSR uses dynamic source routing [13] and it adapts quickly to routing changes when host movement is frequent, however it requires little or no overhead during periods in which host moves less frequently. Source routing is a routing technique in which the sender of a packet determines the complete sequence of nodes through which to forward the packets, the sender explicitly lists this route in the packet’s header, identifying each forwarding hop by the address of the next node to which to transmit the packet on its way to the destination host. The protocol is designed for use in the wireless environment of an ad hoc network. Route cache is maintained to reduce cost of route discovery. Route Maintenance is used when sender detects change in topology or source code has got some error. In case of errors sender can use another route or invoke Route Discovery again. The DSR is single path routing.

Temporary Ordered Routing Algorithm (TORA)

TORA [14] uses the Link reversal technology. It is structured as a temporally ordered sequence of diffusing computations; each computation consisting of a sequence of directed link reversals. It is based on LRR (Link reversal routing). The protocol is highly adaptive, efficient, loop free and scalable. Important concept in its design is that it decouples the generation of potentially far-reaching control message propagation from the rate of topological changes. It reduces energy consumption without diminishing the capacity or connectivity of the network.

Ad hoc On demand Distance Vector (AODV)

Ad Hoc On Demand Distance Vector (AODV) [15] is pure on demand routing system The AODV routing protocol is intended for use by mobile nodes in an ad hoc network characterized by frequent changes in link

connectivity to each other caused by relative movement. It offers quick adaptation to dynamic link conditions, low processing and memory overhead, low network utilization, and establishment of routes between sources and destination which is loop free at all times. It follows quick adaptation to changes. It has low memory overhead.

4.0 Proposed Scheme

The objective of the proposed self healing scheme is to design a scalable routing protocol for large scale networks. It uses concept of Backbone nodes network.

4.1 Mobile Backbone Networks (MBNs): A backbone network is a network consisting of a large area with hundreds of nodes. There are two types of nodes in these networks: backbone nodes and regular nodes (RNs). Since the BNs are also mobile and keep joining and leaving the backbone network in an ad hoc manner, the backbone network is actually a MANET. Thus, there are multiple MANETs in a multi-level MBN. All nodes in a network operate in the same channel but these networks operate in different channels to minimize the interference across levels. There are three critical issues involved in building a MBN as

1. Number of backbone Nodes
2. Deployment of backbone Nodes
3. Routing Protocols

4.1.1 Number of Backbone Nodes

Optimal number of backbone nodes has been calculated with the aim of maximizing per node throughput. In general, the network is designed such that it has sufficient number of BNs to cover the whole network.

4.1.2 Deployment of backbone Nodes

Ideally, backbone node (BNs) should be deployed such that the number of BNs to cover the whole network is optimal. This could be done by pre-assigning BNs and scattering them around the terrain at the time of network initialization. However, this may not be worth because these BNs are also moving, and may go down which may leave some Routing Nodes having no BNs to associate with. The typical solution is to deploy redundant Nodes in the network and elect some of them as BNs. The task of selecting BNs from network is called backbone election. When all BNs move out of the reach of a network then that network changes its status. Hence, management of number and deployment of BNs are completely distributed, dynamic and self-organized. It is desired to perform in a distributed manner and dynamically in such a manner that the BNs are scattered in the terrain.

4.1.4 Routing Protocols

After a set of BNs is elected and these BNs are connected through a high power radio to form the backbone network, the one critical issue remained is routing. There are ample choices available to select a routing protocol. Apart from the application, the one of the most important consideration while choosing a routing protocol is that it should be able to utilize the short cut and additional bandwidth provided by the separate high power links among BNs. AODV has been used as base protocol and changes have been made to it.

In this paper a new scheme, known as the Backbone nodes network [16] has been suggested which would allow mobile nodes to maintain routes to destinations with more stable route selection. This scheme responds to link breakages and changes in network topology in a timely manner. It uses concept of backbone nodes network as explained earlier. This makes route maintenance and recovery phase more efficient and fast. This backbone nodes network helps in reconstruction phase in the fast selection of new routes. Each route table has an entry for number of backbone nodes attached to it. Whenever need for a new route arises in case of route break, check for backbone nodes are made, and a new route is established. Same process is repeated in route repair phase. Route tables are updated at each hello interval as in AODV with added entries for backbone nodes. These are nodes at the one hop distance from its neighbor. Backbone nodes are those nodes which are not participating in route process currently or nodes which enter the range of transmission during routing process. As nodes are in random motion for a scenario, so there is every possibility that some nodes are idle and are in the vicinity of the routing nodes. Whenever a break in the route phase occurs due to movement of participant node, node damage or for other reasons; these idle nodes which have been termed as backbone nodes take care of the process and start routing. The whole process becomes fast and more packet delivery is assured. The changes in the existing protocol are required at route reply and route recovery phases. In these phases the route table is updated with entries of backbone nodes. Each route table has an entry for number of backbone nodes surrounding it and their hop distance from the node. For simplicity of the protocol the distance has been assumed to be one hop.

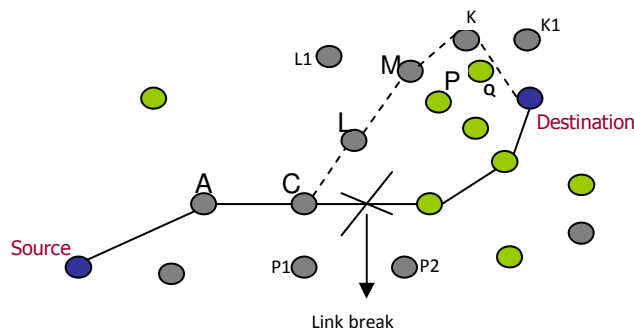


Figure 2: Local repair

Figure 2 gives an idea of self healing process. Initial path from source node 'Source' to destination node 'Destination' is shown via solid lines. When link breaks at node C, route repair healing starts, node C starts searching for new paths. Node C invokes Route Request phase for 'Destination'. Now backbone nodes are selected and proper selection of nodes is done. Path selected becomes [C - L - M - K - Destination]. If any BN has not been on active scale, it is rejected and a new node is searched. In addition to power factor, efforts are made to keep the path shortest. This healing process attempts are often invisible to the originating node.

5.0 Conclusion

In this paper a new scheme has been presented that utilizes backbone network. The scheme can be incorporated into any ad hoc on-demand protocol to heal link failures. It will improve reliable packet delivery even in route breaks. As a case study, the proposed scheme has been incorporated to AODV and it is expected that the performance improves. Study is going on currently investigating ways to make this new scheme robust to traffic load. The proposed scheme gives a better approach for on demand routing protocols for route selection and maintenance. It is expected that overhead in this protocol will be slightly higher than others, which is due to the reason that it requires more calculations initially for checking backbone nodes. This also may cause a bit more end to end delay. The proposal is to check this scheme for more detailed and realistic channel models with fading and obstacles in the simulation. Efforts are on to simulate the scheme using NS2 and compare results with existing schemes.

Self-healing systems are relatively new both for the academia and the industry. However, hope is to see a large number of systems, software and architectures that borrow from nature, ideas and concepts very quickly in future. Modeling computer security using biology as a motivation can help in creating adaptive systems that provide functionality despite the possibility of disasters. The obvious goal is to generate a technique that will reveal that Self-healing networks are designed to be robust even in environments where individual links are unreliable, making them ideal for dealing with unexpected circumstances. The dynamic nature that gives these networks their self-healing properties, however, also makes them difficult to test. Even after multiple deployments and thorough simulation, it's difficult to predict how these systems will work (or fail) in actual emergencies. Though the best uses for technologies are often difficult to predict, still one can almost certain that the self-healing networks is waiting to be developed and getting popular.

REFERENCES

- [1] S. H. Bae, S. J. Lee, W. Su, and M. Gerla, "The Design, Implementation, and Performance Evaluation of the On-Demand Multicast Routing Protocol in Multihop Wireless Networks", *IEEE Network, Special Issue on Multicasting Empowering the Next Generation Internet*, vol. 14, no. 1, January/February 2000.
- [2] Glab, M., Lukasiewicz, M., Streichert, T., Haubelt, C., Teich, J.: Reliability-Aware System Synthesis. In: Proceedings of DATE 2007, pp. 409–414 (2007).
- [3] D. Bertsekas and R. Gallager, "Data Networks" Prentice Hall Publ., New Jersey, 2004.

- [4] Robert Poor, Cliff Bowman, Charlotte Burgess Auburn, 'Self healing networks' ACM Queue, pp 52-59, 2003
- [5] A.Kush, R.Chauhan, P.Gupta, "Power Aware Virtual Node Routing Protocol for Ad hoc Networks" in International Journal of Ubiquitous Computing and Communication (UBICC), Vol2 No. 3 , pp 56-61, South Korea 2008.
- [6] Hiroyuki Fujii and Noriaki Yoshikai. 'Restoration message transfer mechanism and restoration characteristics of double-search self-healing atm network'. IEEE Journal on Selected Areas in Communications, 12(1): pp 149-158, Jan 2004.
- [7] Ryutaro Kawamura, Kenichi Sato, and Ikuo Tokizawa. Self-healing atm networks based on virtual path concept. IEEE Trans. Communications, 12(1):120-127, Jan 2004.
- [8] K. Murakami and H. Kim. Optimal capacity and flow-assignment for self-healing atm networks based on line and end-to-end restoration. IEEE/ACM Transactions on Networking, 6(2):207- 221, Apr 2008.
- [9] S. Kwong, H.W. Chong , M. Chan, K.F. Man 'The use of multiple objective genetic algorithm in self-healing network' September 2002, Elsevier Science B.V.
- [10] Cankay, H.C., Nair, V.S.S.: Accelerated reliability analysis for self-healing sonet networks.SIGCOMM Computer Communications. Rev. 28(4), pp 268–277, 1998.
- [11] Robert Poor, Cliff Bowman, Charlotte Burgess Auburn, 'Self healing networks' ACM Queue, pp 52-59, 2003
- [12] S.W. Cheng, D. Garlan, B. Schmerl, P. Steenkiste, N. Hu, "Software architecture-based adaptation for grid computing", The 11th IEEE Conference on High Performance Distributed Computing (HPDC'02), Edinburgh, Scotland., 2002.
- [13] D. B. Johnson et al., "The dynamic source routing protocol for mobile ad hoc networks (DSR)", Internet Draft, MANET working group, Feb 2002.
- [14] V. Park, S. Corson, "Temporally-Ordered Routing Protocol (TORA) Specification", Internet Draft, October 1999.
- [15] C. Parkins and E. Royer, "Ad Hoc on demand distance vector routing", 2nd IEEE workshop on mobile computing , pages 90-100, 1999
- [16] A.Kush, Divya S., "Power Aware Routing Scheme in Ad Hoc Net" IJCSI International Journal of computer Science Issues, Vol. 7, Issue 1, 2010, ISSN (Online): 1694-0784, ISSN (Print): 1694-0814.

Vectorization Algorithm for Line Drawing and Gap filling of Maps

Ms. Neeti Daryal
*Lecturer, Department of Computer Science,
M L N College, Yamuna Nagar*

Dr Vinod Kumar
*Reader, Department of Mathematics
J.V. Jain College, Saharanpur*

Abstract

Vectorization, i.e. raster-to-vector conversion is heart of graphics recognition problems, as it deals with converting the scanned image to a vector form suitable for further analysis. Many vectorization methods have been designed. This paper deals with the method of raster-to-vector conversion which proposed for capturing line drawing images. In the earliest works on vectorization, only one kind of method was introduced. The proposed algorithm combines the features of thinning method and medial line extraction method so as to produce best line fitting algorithm. There are several steps in this process. The first step is Pre-processing, in which find the line into original raster image. Second is developing an algorithm for gap filling between the adjacent lines to produce vectorization for scanned map. Result and Literature about the above mentioned methods is also included in this paper.

Key Words: Vectorization, Gap filling, Line drawing, Thinning algorithm, Medial algorithm

1. INTRODUCTION

Graphics recognition is concerned with the analysis of graphics-intensive documents, such as technical drawings, maps or schemas. Vectorization, i.e. raster-to-vector conversion, is of course a central part of graphics recognition problems, as it deals with converting the scanned map to a vector form suitable for further analysis. Line drawing management systems store visual objects as graphic entities. Many techniques have already been proposed for the extraction and recognition of graphic entities from scanned binary maps. In particular, various raster-to-vector conversion methods have been developed which convert image lines

into vector lines automatically. In this paper, a new raster-to-vector conversion method is proposed for capturing high-quality vectors in a line drawing.

Bitmap Image:

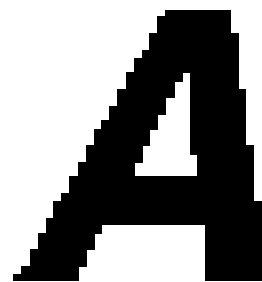


Figure 1[1]: Raster Vector Graphics

Vector Graphic:

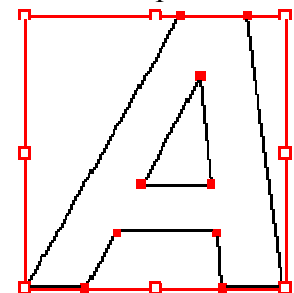


Figure1[2] Graphics

There are two kinds of computer graphics - raster (composed of pixels)

and vector (composed of paths)[1]. Raster images are more commonly called bitmap images. Vector graphics are called object-oriented graphics as shown in Figure 1[2].

2. NEED OF VECTORIZATION

In general, vector data structure produces smaller file size than raster image because a raster image needs space for all pixels while only point coordinates are stored in vector representation [3]. This is even truer in the case when the graphics or images have large homogenous regions and the boundaries and shapes are the primary interest.

3. RELATED WORK

Vectorization techniques have been developed in various domains and a number of methods have been proposed and implemented. These methods are roughly divided into two classes: Thinning based methods and Non-thinning based methods [4].

Thinning based methods are applied in most of the earlier vectorization schemes [4]. These methods usually employ an iterative boundary erosion process to remove outer pixels until only one-pixel-wide skeleton remains like "peeling an onion" [5]. A polygonal approximation procedure is then applied to convert the skeleton to a vector, which may be a line segment or a plotline. The thinning method tends to create noisy junctions at corners, intersections, and branches as shown in the Figure 2[6].

Among the non-thinning based methods. Medial line extraction methods, surveyed in were also popular in the early days of vectorization [7]. Methods of this class extract image contours (the

edges of the shape) before the medial axis between the two side edges is

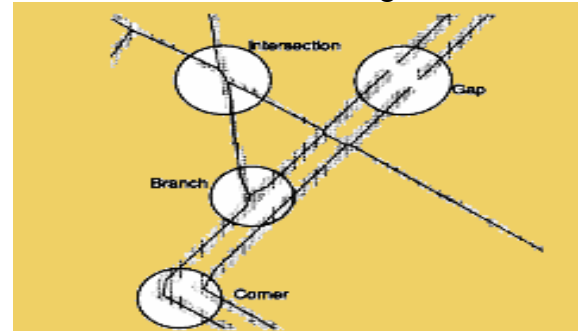


Figure 2: Defects of the thinning method

found. The midpoint of two parallel lines is given by the midpoint of a perpendicular line projected from one side to the other, and these midpoints are coordinates which represent vectors [5]. The medial line extraction method often misses pairs of contour lines at branches as shown in Figure 3[6] consequently it fails to find the midpoint of parallel lines [8].

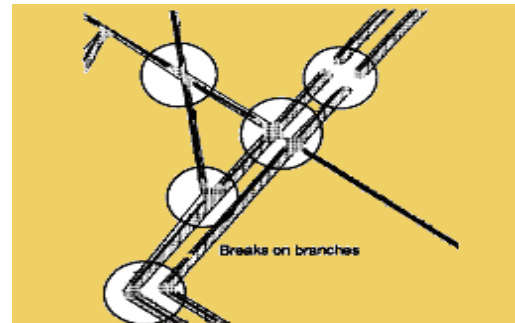


Figure 3: Defects of the medial line extraction method

Other classes of non-thinning based methods that also preserve line width have been developed recently [5]. These include run graph based methods mesh pattern based methods and the Orthogonal Zig-Zag (OZZ) method. These methods are not included in this paper. We are working with above said two methods only.

The disadvantages of thinning based methods and medial line extraction

methods lead to a failure in fitting a line properly. But the thinning method is able to maintain connectivity but loses shape information. Interestingly, the medial line extraction method has the complementary features; that is, it maintains shape information but tends to lose line connectivity. In combination, they could be realized; good-quality extracted lines could be obtained.

4. PROPOSED VECTORIZATION PROCESS

The following is an implementation of the line fitting concept. The purpose of the particular method has been carefully designed to offer practical performance with both acceptable processing speed and good vector quality. Figure 4 shows a flowchart for the whole procedure [5].

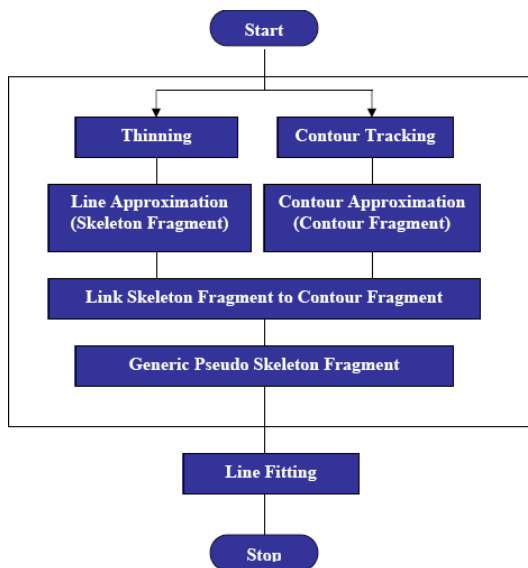


Figure 4: Flowchart of line fitting method based on contours and skeletons

Basic vectorization requires the following tasks:

(1) Linking short line Segments into longer integrated ones.

(2) Correcting the defects at junctions.

(3) Modifying vector attributes such as endpoints intermediate points and line width.

Linking short line segments into longer ones may yield the correct line width and overcome some junction problems. Other defects at junctions, such as corners and branches are subject to special processing [9]. The precise intersection points. i.e. the endpoints of the vectors, are calculated.

The combination has several steps in this process.

The first step is *preprocessing* in which find the line into original raster image. Second is Gap filling between the adjacent lines.

4.1 PREPROCESSING

- A scanned line drawing is converted from binary raster image data to run length code data.
- Processed into skeletons and Tracked for contours.
- Each skeleton fragment is linked to neighboring contour fragments.
- Processed into skeleton and contour fragment respectively.

4.2 GAP FILLING

In a contour image the contour lines are split and the different contour levels are written in the gap. This causes problems in automatic vectorization of images. Since the text are erased and not taken into account while vectorizing, the final

output has gaps in between lines. Gaps are also produced due to noise. Thus gap filling should be given prime importance after processed into skeleton and contour fragment respectively. A poor-quality line drawing often has gaps which prevent correct vector extraction [10]. Following algorithm shows the steps for gap filling

ALGORITHM

Step 1: Reading the input and getting the x and y coordinates of the line.

Step2: Get x and y coordinates of the endpoints.

Step3: Find distance between endpoints. After finding the end points we find the distance between the end points using the Euclidean distance formula which can be mathematically represented as,
$$D = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$
Where D is the Euclidean distance and (x_1, y_1) and (x_2, y_2) are endpoints.

Step 4: IF distance < threshold then set the threshold otherwise stop.

Step5: Setting the threshold.

Step6: Get the x and y coordinate of end points and five adjacent points corresponding to the line then we get the x and y coordinate of the end points that have distance that is less than the set threshold.

Step7: Check if any of the distance are equal then we go to step 8 (slope function) otherwise go to step 9 (Least Square Parabola).

Step8: Slope Function

Step9: Least Square Parabola fitting.

Using these coordinates we perform least square parabola fitting to get the values of the coefficient, a, b and c. Using the values of a, b and c and the x coordinates of the two lines we can get an approximate value of y. There are other cases where we can directly extend the line and we do not have to approximate the curve. The X and Y coordinate are chosen based on four cases as shown below. Let us consider that (x_1, y_1) and (x_2, y_2) are the end points of two lines whose distance is less than the threshold value.

Consider Figure 6, the end points are highlighted in red. Here we can see that $x_1 \neq x_2$ and $y_1 \neq y_2$ and $x_1 \neq y_1$ and $x_2 \neq y_2$. In this case since $x \neq y$ we cannot connect it using a straight line and so we will use Least Square parabola to interpolate the points in between the endpoints. Using the x and y coordinates of the two lines we get the value of a, b and c using the steps explained in the above section. After we get the values of a, b and c we increment minimum value of x by 1 until it reaches maximum value of x and substitute the value of x in the following equation to get the corresponding y value,

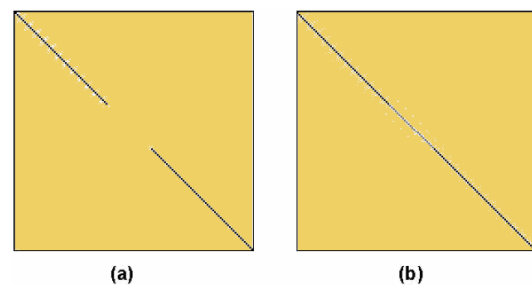


Figure 5: Example of Gapfilling

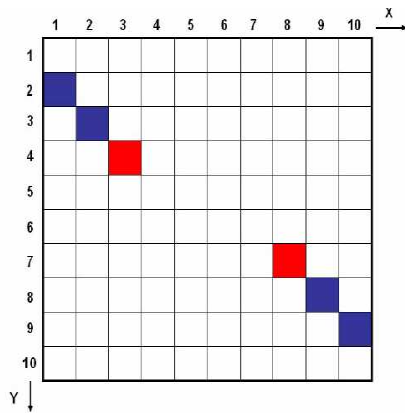


Figure 6: Case 4: Gap Filling
 $f(x) = a + bx + cx^2$

Where $f(x) = y$. The least squares line method uses this equation to get the parabola graph. After getting the value of y we approximate the number to a natural number. The condition for approximation being that if the decimal value is greater than or equal to 0.5 then it is approximated (rounded) to the next number and if it is less than 5 then it is approximated to the real number. For example if the value of y is 4.75 then it is approximated to 5 and if the value of $y = 4.30$ then it is approximated to 4. An example of gapfilling of this case is shown in Figure 7.

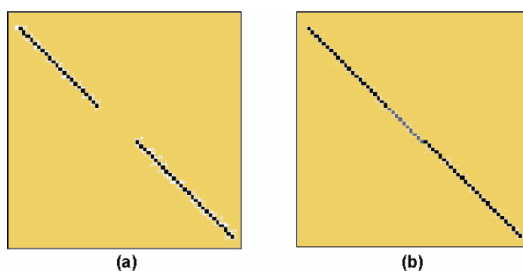


Figure 7: Example of Gapfilling

Rounding the number or approximating is only done for raster images and not for vector data since there is no need to rasterized the curve. LSP is used only for case four because in the other cases we get the exact coordinates by just extending the line and we do not have to

interpolate and get the x coordinate to get the corresponding y coordinate.

5. Results

The result obtained has been shown using all the foresaid discussed methods displayed in the form of results as follows. Figure 8 is the scanned image and Figure 9 the corresponding gap filled image. Since this is an iterative process all the gaps that are within the threshold are filled.

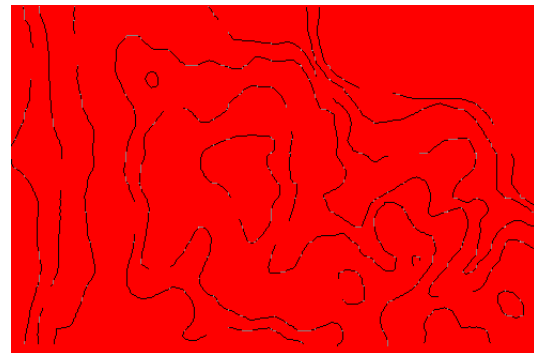


Figure 8: Contour Image with Gap

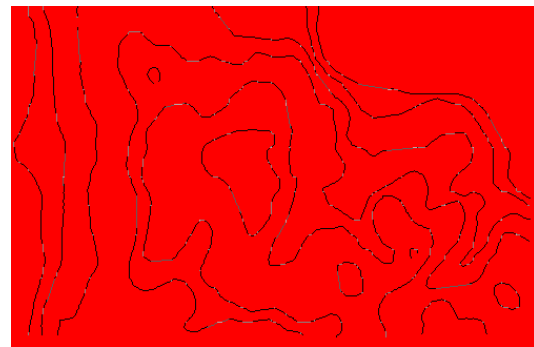


Figure 9: Gap Filled Contour Image

6. CONCLUSION

In this paper, we have discussed the line formation, which has been done through the combination of line fragment and contour fragment algorithm for building a vectorization method which leads to filling the gap between the lines. More specifically, the gap between the lines have been filled by Least Square

Parabola fitting algorithm This resultant of this method has been applied for the correction of scanned map, shown as Figure 8 & 9.

6. REFERENCES

- [1] J.Jimenez and J .L.Navalon, "Some experiments in image vectorization," IBM J. Res. Develop. 26, pp.724-734(1982) [4] R.O.Duda, P.E.Hart, "Use of Hough transformation to detect lines and curves in pictures," Commun.ACM, 15, 1, pp.11-15(1972) [5] J. Jimenez and J.L. Navalon, -'Some Experiments in Image Vectorization' , IBM J. Res. Develop. 26, pp724-734, 1982.
- [2] Smith R.W. (1978). Computer processing of line images: A survey. Pattern Recognition x; 20(1):7-15.
- [3] R.Kasturi, S.Siva, and L.O'Gorman, "Techniques for Line Drawing Interpretation: An Overview," Proc. IAPR Workshop on Machine Vision Applications, pp. 15 1-160(1990)
- [4] H.Tamura, "A Comparison of line thinning algorithms from digital geometry viewpoint," Proc.4th Int. Jt Conf. on Pattern Recognition, Kyoto, Japan, pp715719, IEEE(1978).
- [5] F.Chang, Y.-C. Lu, and T. Pavlidis. Feature Analysis Using Line Weep Thinning Algorithm. IEEE Transactions on PAMI, 21(2):145-158, Feb. 1999.
- [6] Tainura, "A Comparison of Line Thinning Algorithms from Digital Geometry Viewpoint", Proc. of 4th hit. Jt. Conf. on Patten Recognition. Kyoto. Japan, pp715-719, 1978.
- [7] Kasturi, S. T. Bow. W. El-Masri. J. Shah, J. R. Gattiker, and U. B. Mokate; "A System for Interpretation of Line Drawings", IEEE Trans. on PAMI, 12(IO), pp978-992, 1990.
- [8] Borgefors. Distance Transforms in Digital Images. Computer Vision, Graphics and Image Processing, 34:344-371, 1986.
- [9] J.Canny. A Computational Approach to Edge Detection. IEEE Transactions on PAMI, 8(6):679-698, 1986.\
- [10] R.W. Smith, "Computer Processing of Line Images: "A Survey", Patteni Recognition, 20(1), pp7-15, 1987.

Simulation Modeling of Reactive Protocols for Adhoc Wireless Network

Sunil Taneja

Department of Computer Science,
Government Post Graduate College,
Kalka, India
suniltaneja.iitd@gmail.com

Ashwani Kush

Department of Computer Science,
University College, Kurukshetra
University, Kurukshetra, India
akush20@gmail.com

Amandeep Makkar

Department of Computer Science,
Arya Girls College,
Ambala Cantt, India
aman.aryacollege@gmail.com

Abstract

Ad hoc wireless networks are characterized by multihop wireless connectivity, infrastructureless environment and frequently changing topology. As the wireless links are highly error prone and can go down frequently due to mobility of nodes, therefore, stable routing is a very critical task due to highly dynamic environment in adhoc wireless networks. In this research paper, simulation modelling of prominent on-demand routing protocols has been done by presenting their functionality using NS2. An effort has been made to evaluate the performance of DSR and AODV on a self created scene using TCL for varying number of mobile nodes. The performance differential parameters analyzed are; packet delivery ratio and sent & received packets with varying speed and pause time. Subsequently, using results obtained after simulation, the recommendations have been made about the significance of either protocol in various situations. It has been observed that both DSR and AODV are good in performance in their own categories but the emphasis for stable and reliable routing is still on AODV as it performs better in denser environments.

Keywords: Adhoc Wireless Networks, DSR, AODV, Routing, Simulation

I. INTRODUCTION TO ADHOC WIRELESS NETWORKS

The wireless networks are classified as Infrastructured or Infrastructure less. In Infrastructured wireless networks, the mobile node can move while communicating, the base stations are fixed and as the node goes out of the range of a base station, it gets into the range of another base station. In Infrastructureless or Ad Hoc wireless network [15], the mobile node can move while communicating, there are no fixed base stations and all the nodes in the network act as routers. The mobile nodes in the Ad Hoc network dynamically establish routing among themselves to form their own network 'on the fly'. In this research paper, intend is to study the mobility patterns of DSR and AODV using simulation modeling by varying 'number of mobile nodes', 'speed', 'pause time', 'UDP connections' and 'TCP connections'. A Mobile Ad Hoc Network is a collection of wireless mobile nodes forming a temporary network without any fixed infrastructure where all nodes are free to move about arbitrarily and where all the nodes configure themselves. Unlike traditional networks whereby routing functions are performed by dedicated nodes or routers, in MANET, routing functions are carried out by all available nodes. There are no fixed base stations and each node acts both as a router and as a host. Even the topology of network may also change rapidly. The mobile nodes in the Ad Hoc network dynamically establish routing among themselves to form their own network 'on the fly'. In essence, the network is created in ad-hoc fashion by the participating nodes without any central administration. Further ad hoc networks

can be classified as single-hop or multi-hop. In single-hop ad hoc networks, nodes are in their reach area and can communicate directly but in case of multi-hop, some nodes are far and cannot communicate directly. The traffic has to be forwarded by other intermediate nodes. Ad hoc networks are primarily meant for use by military forces or for emergency rescue situations. At the state of war an army cannot rely on fixed infrastructure, as it is an easy and attractive target for the enemy. Ad hoc networks are optimal solution in such cases. For civil use ad hoc networks are crucial if the fixed infrastructure has been torn down by some natural disaster, like a flood or an earthquake. Then rescue operations could in such a situation be managed through utilizing ad hoc networks. Mobile ad hoc networks have several advantages over traditional wireless networks including ease of deployment, speed of deployment and decreased dependence on a fixed infrastructure but there are certain open research issues too in its implementation. Some of the research issues [15] are: *Dynamic topology, Autonomous or no centralized administration, Device discovery, Bandwidth optimization, Scalability, Limited security, Power Aware Routing, Self healing, Poor transmission quality, Ad hoc addressing, and Topology maintenance.*

In Ad hoc network, nodes can change position quite frequently. Each node participating in the network must be willing to transfer packets to other nodes. For this purpose, a routing protocol is needed. Our focus in this research paper is on the stable and reliable routing over mobile adhoc networks. The proposed routing scheme is to select a stable and reliable path in such a manner that load is balanced over the entire network.

II. ROUTING PROTOCOLS

A routing protocol [15] is required whenever a packet needs to be transmitted to a destination via number of nodes and numerous routing protocols have been proposed for such kind of ad hoc networks. These protocols find a route for packet delivery and deliver the packet to the correct destination. The studies on various aspects of routing protocols [1, 2] have been an active area of research for many years. Many protocols have been suggested keeping applications and type of network in view. Basically, routing protocols can be broadly classified into two types as: Table Driven Protocols or Proactive Protocols and On-Demand Protocols or Reactive Protocols. In Table Driven routing protocols each node maintains one or more tables containing routing information to

every other node in the network. All nodes keep on updating these tables to maintain latest view of the network. Some of the existing table driven protocols are DSDV [4], GSR [9], WRP [8] and ZRP [11]. In on-demand routing protocols, routes are created as and when required. When a transmission occurs from source to destination, it invokes the route discovery procedure. The route remains valid till destination is achieved or until the route is no longer needed. Some of the existing on demand routing protocols are: DSR [5], AODV [3] and TORA [10]. The emphasis in this research paper is concentrated on the study of mobility pattern and performance analysis of two prominent on-demand routing Protocols i.e. DSR and AODV. Surveys of routing protocols for ad hoc networks have been discussed in [12, 13, 14]. A brief review of DSR and AODV is presented here as these have been analyzed in this paper for their performance.

(a) **DSR** [5, 7] is an Ad Hoc routing protocol which is source-initiated rather than hop-by-hop and is based on the theory of source-based routing rather than table-based. This is particularly designed for use in multi hop wireless ad hoc networks of mobile nodes. Basically, DSR protocol does not need any existing network infrastructure or administration and this allows the Network to be completely self-organizing and self-configuring. This Protocol is composed of two essential parts of route discovery and route maintenance. Every node maintains a cache to store recently discovered paths. When a node desires to send a packet to some node, it first checks its entry in the cache. If it is there, then it uses that path to transmit the packet and also attach its source address on the packet. If it is not there in the cache or the entry in cache is expired (because of long time idle), the sender broadcasts a route request packet to all of its neighbors asking for a path to the destination. The sender will be waiting till the route is discovered. During waiting time, the sender can perform other tasks such as sending/forwarding other packets. As the route request packet arrives to any of the nodes, they check from their neighbor or from their caches whether the destination asked is known or unknown. If route information is known, they send back a route reply packet to the destination otherwise they broadcast the same route request packet. When the route is discovered, the required packets will be transmitted by the sender on the discovered route. Also an entry in the cache will be inserted for the future use. The node will also maintain the age information of the entry so as to know whether the cache is fresh or not. When a data packet is received by any intermediate node, it first checks whether the packet is meant for itself or not. If it is meant for itself (i.e. the intermediate node is the destination), the packet is received otherwise the same will be forwarded using the path attached on the data packet. Since in Ad hoc network, any link might fail anytime. Therefore, route maintenance process will constantly monitors and will also notify the nodes if there is any failure in the path.

Consequently, the nodes will change the entries of their route cache.

(b) **AODV** [3, 7] is a variation of Destination-Sequenced Distance-Vector (DSDV) routing protocol which is collectively based on DSDV and DSR. It aims to minimize the requirement of system-wide broadcasts to its extreme. It does not maintain routes from every node to every other node in the network rather they are discovered as and when needed & are maintained only as long as they are required. The algorithm used by AODV for establishment of unicast routes can be summarized as. When a node wants to send a data packet to a destination node, the entries in route table are checked to ensure whether there is a current route to that destination node or not. If it is there, the data packet is forwarded to the appropriate next hop toward the destination. If it is not there, the route discovery process is initiated. AODV initiates a route discovery process using Route Request (RREQ) and Route Reply (RREP). The source node will create a RREQ packet containing its IP address, its current sequence number, the destination's IP address, the destination's last sequence number and broadcast ID. The broadcast ID is incremented each time the source node initiates RREQ. Basically, the sequence numbers are used to determine the timeliness of each data packet and the broadcast ID & the IP address together form a unique identifier for RREQ so as to uniquely identify each request. The requests are sent using RREQ message and the information in connection with creation of a route is sent back in RREP message. The source node broadcasts the RREQ packet to its neighbours and then sets a timer to wait for a reply. To process the RREQ, the node sets up a reverse route entry for the source node in its route table. This helps to know how to forward a RREP to the source. Basically a lifetime is associated with the reverse route entry and if this entry is not used within this lifetime, the route information is deleted. If the RREQ is lost during transmission, the source node is allowed to broadcast again using route discovery mechanism. Maintenance of routes is done using Local route repair scheme.

III. COMPARATIVE STUDY OF DSR AND AODV

DSR and AODV share certain salient characteristics. Specifically, they both discover routes only in the presence of data packets in the need for a route to a destination. Route discovery in either protocol is based on query and reply cycles and route information is stored in all intermediate nodes on the route in the form of route table entries (AODV) or in route caches (DSR). However, there are several important differences [7, 16] in the dynamics of these two protocols, which may give rise to significant performance differentials. The important differences are given below in the form of benefits and limitations of these protocols. These differences help in studying the pattern analysis and performance evaluation of either protocol.

Benefits and Limitations of DSR

DSR protocol has number of benefits. It does not use periodic routing messages (e.g. no router advertisements and

no link-level neighbor status messages), thereby reducing network bandwidth overhead, conserving battery power, and avoiding the propagation of potentially large routing updates throughout the ad hoc network. There is no need to keep routing table so as to route a given data packet as the entire route is contained in the packet header. The routes are maintained only between nodes that need to communicate. This reduces overhead of route maintenance. Route caching can further reduce route discovery overhead. A single route discovery may yield many routes to the destination, due to intermediate nodes replying from local caches. The DSR protocol guarantees loop-free routing and very rapid recovery when routes in the network change. It is able to adapt quickly to changes such as host movement, yet requires no routing protocol overhead during periods in which no such changes occur. In addition, DSR has been designed to compute correct routes in the presence of asymmetric (uni-directional) links. In wireless networks, links may at times operate asymmetrically due to sources of interference, differing radio or antenna capabilities, or the intentional use of asymmetric communication technology such as satellites. Due to the existence of asymmetric links, traditional link-state or distance vector protocols may compute routes that do not work. DSR, however, will find a correct route even in the presence of asymmetric links.

DSR protocol is not totally free from limitations as it is not scalable to large networks. It is mainly efficient for mobile ad hoc networks with less than two hundred nodes. DSR requires significantly more processing resources than most other protocols. In order to obtain the routing information, each node must spend lot of time to process any control data it receives, even if it is not the intended recipient. The contention is increased if too many route replies come back due to nodes replying using their local cache. The Route Reply Storm problem is there. An intermediate node may send Route Reply using a stale cached route, thus polluting other caches. This problem can be eased if some mechanism to purge (potentially) invalid cached routes is incorporated. The Route Maintenance protocol does not locally repair a broken link. The broken link is only communicated to the initiator. Packet header size grows with route length due to source routing. Flood of route requests may potentially reach all nodes in the network. Care must be taken to avoid collisions between route requests propagated by neighboring nodes.

Benefits and Limitations of AODV

AODV protocol has number of benefits. The routes are established on demand and destination sequence numbers are used to find the latest route to the destination. The connection setup delay is lower. It also responds very quickly to the topological changes that affects the active routes. It does not put any additional overheads on data packets as it does not make use of source routing. It favors the least congested route instead of the shortest route and it also supports both unicast and multicast packet transmissions even for nodes in constant movement.

AODV has also certain limitations like DSR. The intermediate nodes can lead to inconsistent routes if the source

sequence number is very old and the intermediate nodes have a higher but not the latest destination sequence number, thereby having stale entries. The various performance metrics begin decreasing as the network size grows. It is vulnerable to various kinds of attacks as it based on the assumption that all nodes must cooperate and without their cooperation no route can be established. The multiple Route Reply packets in response to a single Route Request packet can lead to heavy control overhead. The periodic beaconing leads to unnecessary bandwidth consumption. It expects/requires that the nodes in the broadcast medium can detect each others' broadcasts. It is also possible that a valid route is expired and the determination of a reasonable expiry time is difficult. The reason behind this is that the nodes are mobile and their sending rates may differ widely and can change dynamically from node to node.

IV. PERFORMANCE METRICS

There are number of qualitative and quantitative performance metrics that can be used to study the mobility pattern of reactive routing protocols viz. packet delivery ratio, average end to end delay, protocol control overhead etc.

Packet Delivery Ratio: This is the ratio of number of packets received at the destination to the number of packets sent from the source. In other words, fraction of successfully received packets, which survive while finding their destination, is called as packet delivery ratio.

Sent and Received Packets: This refers to the number of packets sent over the network by the source node and the number of packets actually received by the destination node.

Average end-to-end delay: This is the average time delay for data packets from the source node to the destination node.

Most of the existing routing protocols ensure the qualitative metrics. Therefore, we have used the packet delivery ratio as quantitative metrics for pattern analysis and performance evaluation of aforementioned routing protocols using simulation modeling for 20 and 50 mobile nodes. The results have also been analyzed for DSR and AODV using sent and received packets with respect to varying speed and pause time.

V. SIMULATION MODEL OF DSR AND AODV

A random waypoint model [17] has been used and some dense/sparse medium scenarios have been generated using TCL. An extensive simulation model having scenario of 20 and 50 mobile nodes is used to study inter-layer interactions and their performance implications. The Simulator used is NS 2.34 [18]. Packet size is 512 bytes. Same scenario has been used for both protocols to match the results. The performance differentials are analyzed using packet delivery ratio with respect to varying speed and pause time and then sent and received packets with respect to speed and pause time.

Packet Delivery Ratio with respect to Speed and Pause Time for simulation of 20 mobile nodes

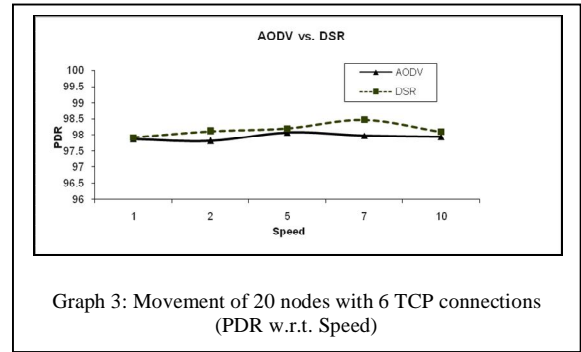
Area considered is 750 meter \times 750 meter and simulation run time is 500 seconds during pattern analysis of 20 nodes using UDP and TCP connections both with respect to varying speed and pause time. Graph 1 shows the packet delivery ratio using speed as a parameter. This performance metric has been evaluated for DSR and AODV using 20 nodes and 6 UDP connections. Speed has been varied from 1m/s to 10 m/s. The

PDR values, computed using received and dropped packets, range from 96% to 99%. The results show that only at one point of time, DSR and AODV gives same PDR value (approx.), otherwise, DSR protocol outperforms AODV in “low mobility” situation.

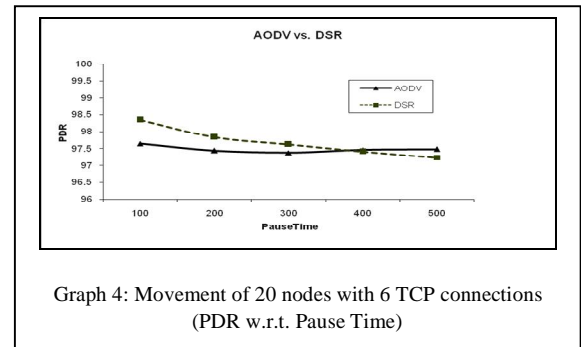
In graph 2, the packet delivery ratio has been evaluated for DSR and AODV protocols using pause time as parameter with same number of nodes and UDP connections. Pause time has been varied from 100s to 500s. The PDR values, computed using received and dropped packets, range from 95% to 99%. In this scenario, the observation is that the DSR protocol outperforms AODV in all the situations.

Graph 3 depicts the packet delivery ratio using speed as a parameter for DSR and AODV protocols. The results are on the basis of 20 mobile nodes and 6 TCP connections. Speed variation is from 1m/s to 10 m/s. The PDR values, computed using received and dropped packets, range from 97% to 99%. The results show that in “low mobility” situation, AODV protocol gives same PDR value (approx.) as that of DSR protocol in the beginning, intermediate and end stage only otherwise, DSR protocol outperforms AODV. On the other hand, AODV outperforms DSR protocol in “high mobility” situation.

In graph 4, the packet delivery ratio has been evaluated using pause time as a parameter on 20 mobile nodes having 6 TCP connections. Pause time varies 100s to 500s. The PDR values, computed using received and dropped packets, range from 97% to 99%. The observation is that the DSR protocol outperforms AODV when pause time is less but AODV outperforms DSR when pause time is high.



Graph 3: Movement of 20 nodes with 6 TCP connections
(PDR w.r.t. Speed)



Graph 4: Movement of 20 nodes with 6 TCP connections
(PDR w.r.t. Pause Time)

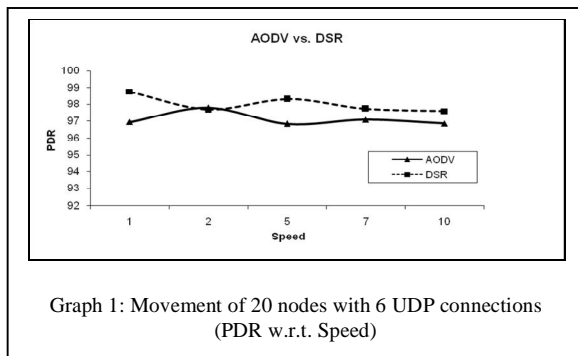
Packet Delivery Ratio with respect to Speed and Pause Time for simulation of 50 mobile nodes

Area considered is 1Km × 1 Km and simulation run time is 700 seconds during pattern analysis of 50 nodes using UDP and TCP connections both with respect to varying speed and pause time. Graph 5 shows the packet delivery ratio using speed as a parameter. This performance metric has been evaluated for DSR and AODV using 50 nodes and 10 UDP connections. Speed has been varied from 1m/s to 10 m/s. The PDR values, computed using received and dropped packets, range from 89% to 95%. The results show that the DSR protocol outperforms AODV.

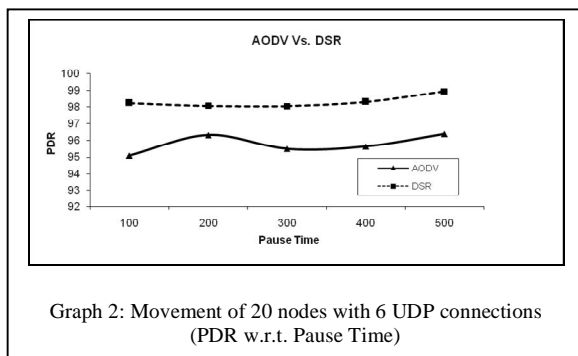
In graph 6, the packet delivery ratio has been evaluated for DSR and AODV protocols using pause time as parameter with same number of nodes and UDP connections. Pause time has been varied from 100s to 650s. The PDR values, computed using received and dropped packets, range from 88% to 95%. In this scenario, the observation is same as above i.e. the DSR protocol outperforms AODV.

Graph 7 depicts the packet delivery ratio using speed as a parameter for DSR and AODV protocols. The results are on the basis of 50 mobile nodes and 10 TCP connections. Speed variation is from 1m/s to 10 m/s. The PDR values, computed using received and dropped packets, range from 91% to 97%. The results show that in “low mobility” situation, AODV protocol gives approximately same PDR value as that of DSR protocol but in “high mobility” situation, AODV outperforms DSR protocol.

In graph 8, the packet delivery ratio has been evaluated using pause time as a parameter on 50 mobile nodes having 10 TCP connections. Pause time varies 100s to 650s. The PDR values, computed using received and dropped packets, range from 92% to 97%. The observation is that the AODV protocol

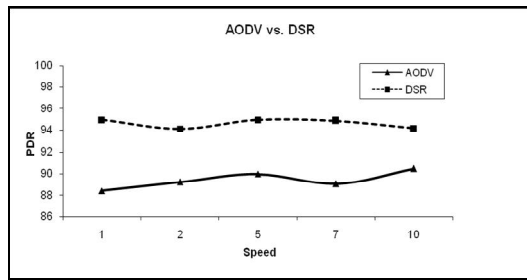


Graph 1: Movement of 20 nodes with 6 UDP connections
(PDR w.r.t. Speed)

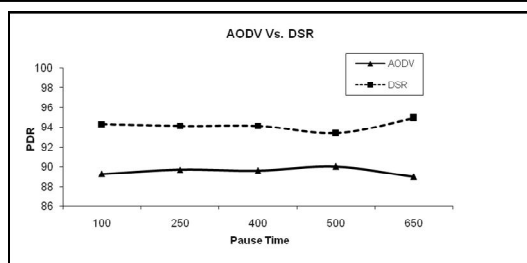


Graph 2: Movement of 20 nodes with 6 UDP connections
(PDR w.r.t. Pause Time)

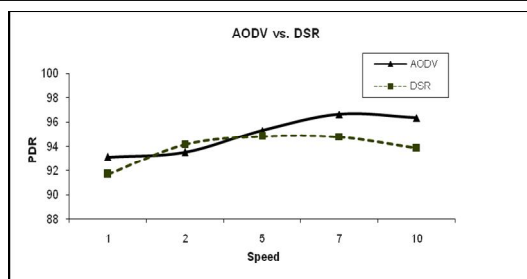
outperforms DSR when pause time is less and AODV protocol gives approximately same PDR value as that of DSR protocol when pause time is high.



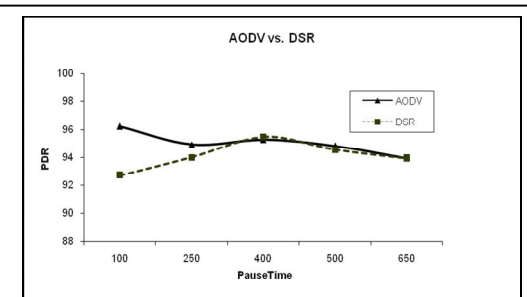
Graph 5: Movement of 50 nodes with 10 UDP connections (PDR w.r.t. Speed)



Graph 6: Movement of 50 nodes with 10 UDP connections (PDR w.r.t. Pause Time)



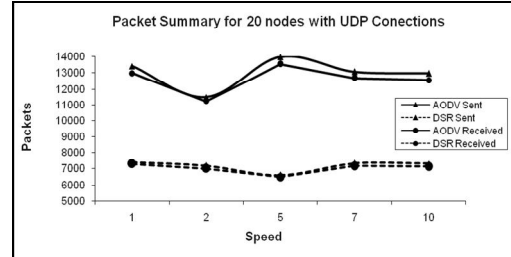
Graph 7: Movement of 50 nodes with 10 TCP connections (PDR w.r.t. Speed)



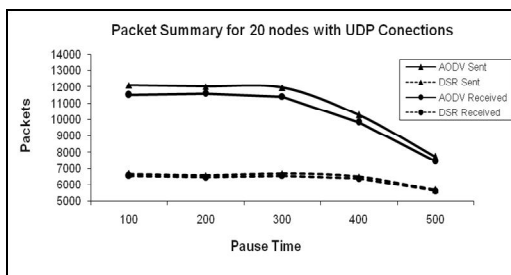
Graph 8: Movement of 50 nodes with 10 TCP connections (PDR w.r.t. Pause Time)

Sent and Received Packets with respect to Speed and Pause Time for simulation of 20 mobile nodes

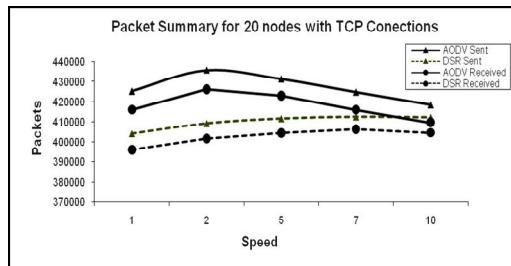
Graph 9 to 12 illustrate the summary of packets sent and received for DSR and AODV protocols with respect to speed and pause time for 20 mobile nodes having 6 UDP and TCP connections.



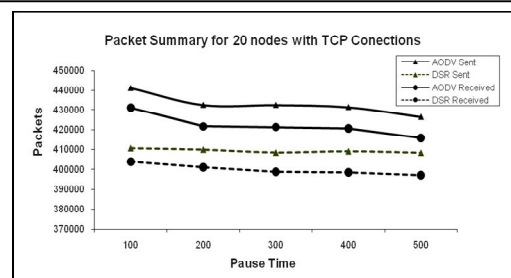
Graph 9: Packets Sent and Received vs. Speed for 20 nodes with UDP Connections



Graph 10: Packets Sent and Received vs. Pause Time for 20 nodes with UDP Connections



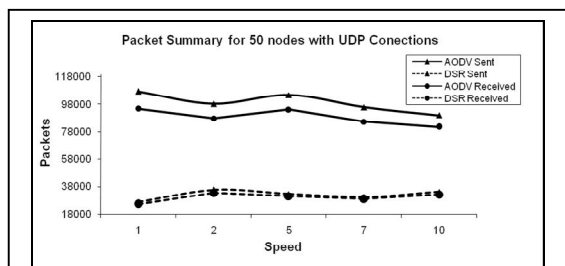
Graph 11: Packets Sent and Received vs. Speed for 20 nodes with TCP Connections



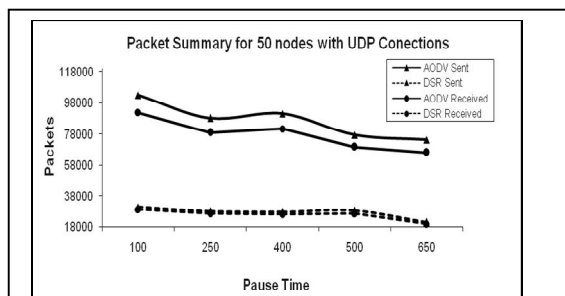
Graph 12: Packets Sent and Received vs. Pause Time for 20 nodes with TCP Connections

Sent and Received Packets with respect to Speed and Pause Time for simulation of 50 mobile nodes

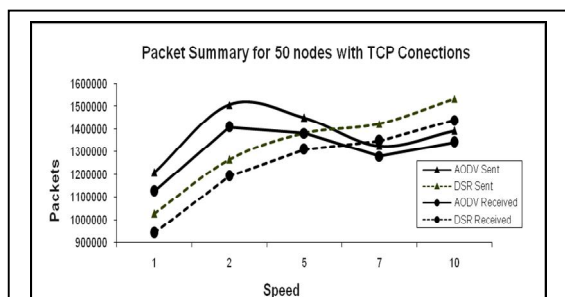
Graph 13 to 16 illustrate the summary of packets sent and received for DSR and AODV protocols with respect to speed and pause time for 50 mobile nodes having 10 UDP and TCP connections.



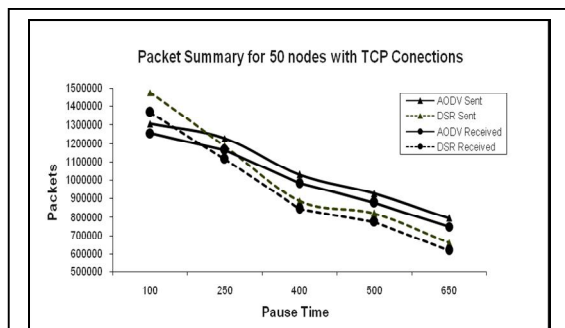
Graph 13: Packets Sent and Received vs. Speed for 50 nodes with UDP Connections



Graph 14: Packets Sent and Received vs. Pause Time for 50 nodes with UDP Connections



Graph 15: Packets Sent and Received vs. Speed for 50 nodes with TCP Connections



Graph 16: Packets Sent and Received vs. Pause Time for 50 nodes with TCP Connections

In this paper, an effort has been made to concentrate on the comparative study and performance analysis of two prominent on demand routing protocols i.e. DSR and AODV on the basis of packet delivery ratio. The earlier work available in the literature has been studied carefully. An effort has been made to perform analysis on a new random way point self created network scenario. The analysis has been reflected in graphs. It has been analyzed that both protocols are good in performance in their own spheres. Still the emphasis of better routing can be on AODV as it performs better in denser mediums. DSR is steady in sparse mediums but it just losses some ground in denser environment and that too when more connections are available and packet are in TCP mode. It is worth mentioning that in the future MANETS, denser mediums will be used with increasing applications, so it can be generalized that AODV is better choice for routing in terms of better packet delivery. Some of the aspects in this study are still under observation as the performance is still to be compared with TORA, STAR and ZRP. More metrics like end to end delay and throughput, load and node life time is still to be taken into account. Power efficiency and secure routing are other major concerns for the future study. An effort will also be made to prove which protocol is best as overall performer.

REFERENCES

- [1] A.Kush, S. Taneja, "A Survey of Routing Protocols in Mobile Adhoc Networks", International Journal of Innovation, Management and Technology, Vol. 1, No. 3, pp 279-285, August 2010.
- [2] A.Kush, R.Chauhan, C.Hwang and P.Gupta, "Stable and Energy Efficient Routing for Mobile Adhoc Networks", Proceedings of the Fifth International Conference on Information Technology: New Generations, ISBN:978-0-7695-3099-4 available at ACM Digital Portal, pp. 1028-1033, 2008.
- [3] C. Perkins, E. B. Royer, S. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing - Internet Draft", RFC 3561, IETF Network Working Group, July 2003.
- [4] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance vector routing (DSDV) for mobile computers", Proceedings of ACM SIGCOMM 94, pp. 34-244, 1994.
- [5] D. B. Johnson, D. A. Maltz, Y.C. Hu, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks (DSR)", IETF Draft, <http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-09.txt>, April 2003.
- [6] P. Chenna Reddy, Dr. P. Chandrasekhar Reddy, "Performance Analysis of Adhoc Network Routing Protocols", Academic Open Internet Journal, Volume 17, 2006.
- [7] Samir R. Das, Charles E. Perkins, Elizabeth M. Royer, "Performance Comparison of Two On-demand Routing Protocols for Ad Hoc Networks", Proceedings of INFOCOM 2000, Tel-Aviv, Israel, March 2000.
- [8] S. Murthy and J. J. Garcia-Luna-Aceves, "An Efficient Routing Protocol for Wireless Networks", ACM Mobile Networks and App. Journal, Special Issue on Routing in Mobile Communication Networks, pp. 183-97, 1996.
- [9] Tsu-Wei Chen and M. Gerla, "Global State Routing: A New Routing Scheme for Ad-hoc Wireless Networks", Proceedings of International Computing Conference IEEE ICC 1998.
- [10] V. Park and S. Corson, Temporally Ordered Routing Algorithm (TORA) Version 1, Functional specification IETF Internet draft, <http://www.ietf.org/internet-drafts/draft-ietf-manet-tora-spec-01.txt>, 1998.
- [11] Z. J. Hass and M. R. Pearlman, "Zone Routing Protocol (ZRP)", Internet draft available at www.ietf.org.
- [12] J. Broch, D. A. Maltz and J. Jetcheva, "A performance Comparison of Multi hop Wireless Adhoc Network Routing Protocols", Proceedings of Mobicomm'98, Texas, 1998.
- [13] S. Ramanathan and M. Steenstrup, "A survey of routing techniques for mobile communications networks", Mobile Networks and Applications, pp. 89-104, 1996.

- [14] E. M. Royer and C. K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks", IEEE Communications, pp. 46–55, 1999.
- [15] C. E. Perkins, Ad Hoc Networking, Addison-Wesley, March 2005.
- [16] S.R. Chaudhry, A. Al-Khwildi; C.Y. Aldelou, H.Al-Raweshidy, "A Performance Comparison of Multi on Demand Routing in Wireless Ad Hoc Networks", IEEE International Conference on Wireless And Mobile Computing, Networking And Communications, Vol. 3, pp. 9 – 16, Aug. 2005.
- [17] Sanchez R., "Mobility Models", at www.disca.upv.es/misan/mobmodel.htm
- [18] NS Notes and Documentation available at www.isi.edu/vint

Media changing the Youth Culture: An Indian Perspective

Prof. Dr. Ritu Soni, Head, Department of Computer Science, Guru Nanak Girls' College,
Yamuna Nagar, Haryana, India-135003

[E-mail- ritusoniynr@yahoo.com](mailto:ritusoniynr@yahoo.com)

Prof. Ms. Bharati Kamboj, Department of Physics, Guru Nanak Girls' College, Yamuna
Nagar, Haryana, India-135003

E-mail- bharatikamboj@gmail.com

Abstract: A study of the effect of electronic media on the Indian youth has been carried out. The internet has revolutionized the entire world. The Indian youth is being exposed to all that is being offered by it. As a result, the Indian traditions are getting diffused to western culture. The way youth is doing purchases or helping others or raising their voices against corruption or any political matters, information access rights are changing with times. Along with the benefits of internet, the youth is also being influenced by its negative aspects. With the technology embedded society emerging, India is undergoing contradicting cultural changes.

Introduction: The term “**culture**” originates from Latin term for “cultivation”, the preparation of land for producing crop[1]. It is defined as all of the beliefs, behaviors, and products common to a particular group. These include the values and customs that we hold in common with others; the language that we speak; the rules we follow; the tools and technologies we use to make things; the goods we consume; the organizations we belong to; and the larger institutions of society[2].

1. Traditional Indian Culture

If we talk about **Indian Culture**, we have certain diverse traditions according to different societies. For example : It is tradition in India that youngsters follow their elders in house hold work & in business, maintain localized traditions for worshipping god, beliefs, cooking traditions, dressing traditions, folk etc. this is possible due to joint family structure available in India. Work is divided between family members, children and older generation get cared, money is pooled, resources are shared. Such a system leads to emotional balance, proper care of every body, high tolerance among family members, helping attitude etc

But due to globalization, the Indian traditions are getting diffused to western culture or within other communities or other social cultures. Media plays a very important role in this kind of diffusion.

How media is changing culture: old media

Different media emerging from time to time- may it be newspaper or radio or films or TV and now internet- has a huge impact on Indian culture as well as society. News paper creates mass awareness, gives right to pen where as radio and films and TV provide entertainment along with mass awareness and right to voice.

With the advent of TV, people are more exposed to different cultures and their traditions. Advertisers use TV as media to promote their products. Social life is taken up by TV rooms. All kinds of programs are viewed by all generations, thereby giving adulthood exposure to children. Thus the cultural impact of all this exposure on society is huge.

Ministry of Broadcasting of India has made some policies for certifying the films but there is no such certification in TV programs.

Amazing internet

And then comes the internet!!!! It has revolutionized the entire world, making it a small global village. Every kind of information is available at the click of mouse. Traditionally, the Internet was used by scientists, academics and engineers to send emails and share information. Today, the internet is also used for the downloading of music, the forming of virtual communities, and the establishment of news groups. **Internet Access Reports** [3] verify that internet is accessed mostly for information retrieval, checking mails and getting sports updates.

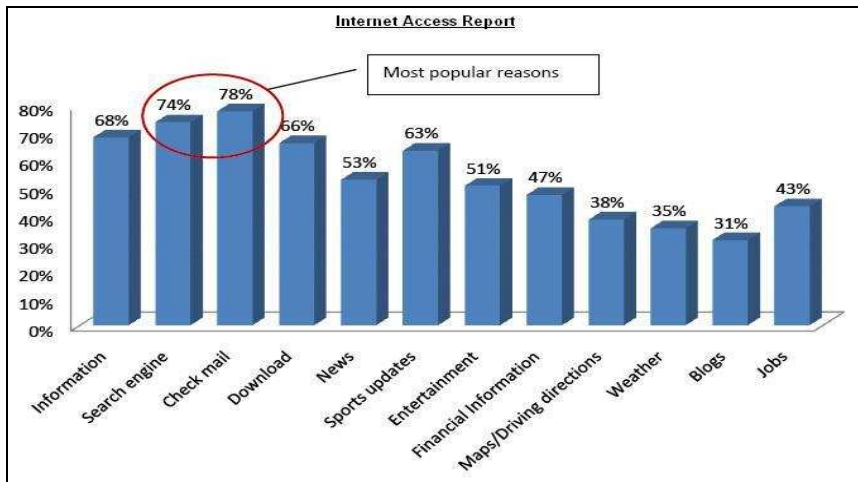


Fig 1

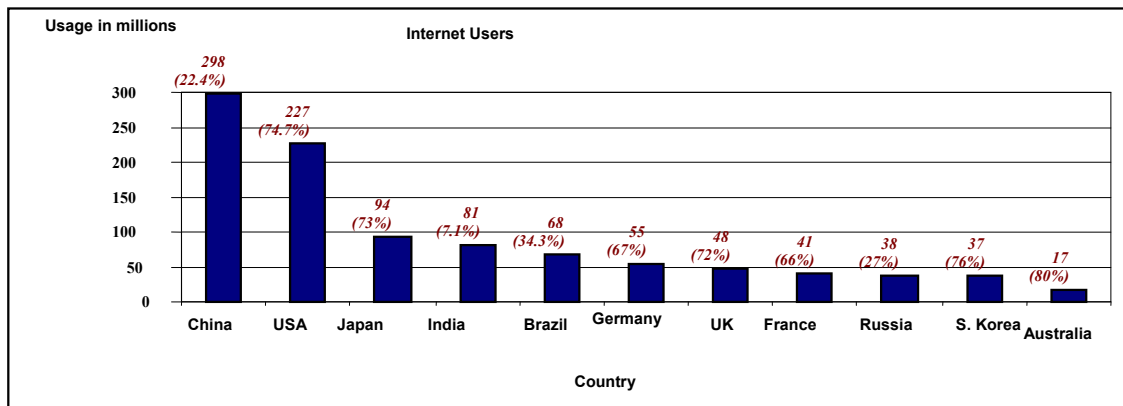


Fig 2

In the age of digital literacy, children could well overtake their parents when it comes to using the internet for social networking, social gathering, tweeting, chatting, sending emails and downloading movies and songs. According to a recent survey by IMRB and IMAI (internet and mobile association of India), nearly 22 million school going children between the ages 8-14 are internet savvy. E-mailing continues to be a popular draw with 91% of those surveyed; general information gathering comes second with 76%.

Nearly 41% of Indians enjoy chatting and online gaming account for 41%. A 2008 study conducted by IMAI showed that 25% of respondent use the internet to hunt for a life partner. (TOI Dec 4th 2009)[4]. Some of the news from Times Of India shows that “India has recently become the third largest market for twitter in just a few months, with 5 million subscribers of total of over 11.5 millions”.

2. Youth population

TCS [5] has identified the youth in four categories -

The Globetrotter: The Globetrotter has global ambitions and wants to study and work abroad. However, a growing confidence in the economic future in India is also reflected in the survey as many students are also keen to bring skills back to India and put them to use here.

The Gadgetphile: Students from both metros and mini-metros who love gadgets and with access to a web enabled mobile, the latest gaming console, i-Pods and if he/she doesn't have one, then aspires to own an i-Phone.

The Nation-Builder: The Indian student is focused on his/her career but is as much interested in the additional benefits that careers brings such as travel, learning new skills, experience to be gained, interesting workplace and salary.

The Social networker: A true digital native, the Social Networker is likely to have as many online friends as real ones and these friendships go beyond the traditional boundaries of gender, caste, and geographies.

History of Youth Culture and New Technologies:

The concept of youth culture and new technologies is a relatively new term. However the rate at which technology has moved into the lives of the young as stated by Seel [7] is unprecedented. Growth within this area began in August 1981 when IBM released the IBM PC [6]. The next major development was the internet established in the 1990's, which “created an entirely new means of social interaction? [7].

More recently, there has been the convergence of individual technologies. An example of this is the convergence of mobile phone technologies with that of the internet.

3. Youth contribution in culture building:

Any culture is reorganized by its unique identity. For youths there are three processes in the formation and construction of their identity; **reflexivity, makeability and individualization** (Aakhus & Katz 2002, p.255). **Reflexivity:** is a cultural way of mirroring (individual, subjective, social, cultural) selves or identities through symbolic images in media, verbal language or face-to-face interaction. **Makeability:** is the perspective that sees personality and lifestyle factors as being open to influence; individuals are able to make choices about their content. **Individualization:** means that as industrial society develops, the significance of social background decreases while the subjective importance of the inner world increases. (Aakhus & Katz, 2002, p.255). The following are some of the areas where the influence of youth culture is recognized:

Youth trends for purchasing:

The effort to commercialize and commodify anything and everything not only leads to the degradation of Indian culture, it is harmful to us as citizens. According to the survey, 57 per cent of the respondents browsed the Internet on their mobile phones, and close to

one-third of those were engaged with brands that advertise. If India could use youth as a hook to lure top brands, it would be the world's most brand-rich country. This is the reason why international brands are making a beeline to tap the Indian youth..

Youth mindset for social networking:

One of the biggest innovations of technology is social networking. A web based social network service focuses on building online communities of people who share interests or activities, attracting people based on common language, sex, religion, or nationality-based identities. These sites also integrate new information and communication tools, such as mobile connectivity, blogging, and photo/video-sharing. Variety of services like creating your own FREE profile webpage, chatting, Instant messaging, playing online games, sharing ideas- sending a friend request or a secret crush etc., are provided by them. It is surveyed that 63% of urban students spend over an hour online daily; 93% are aware of social networking; Orkut and Facebook are most popular online destinations. Other popular social networking sites are MySpace, Friendster, Cyworld, Orkut, Zorpia, Hi5 and Bebo.

Youth mindset for crime:

Every technology has its own negative aspects. **Cyber bullying** is a disturbing new trend among young people abusing the technology. The study shows that 15% of children experience **Cyber sexual harassment** on social sites & chat rooms.

Networking sites also helps burglars home in as they get information via scraps, tweets sent by user to inform their friends that users are away on holiday

Terrorists also use internet to pass their information or scaring the innocent people by sending emails , sms or tweets or by writing blogs. Since media is a one sided communication from “media to masses”; sometime wrong information or view point is also transmitted to masses.

Youth mindset for helping others:

Indians are very helpful by nature. One can find lots of emails asking help for people financially, blood requirements or helping people **in natural disasters** in India.

Colleges are embracing student blogs on their websites, seeing them as powerful tool for admissions. **On-line helps** during admission days are very common in India.

Online car sharing portals are already a popular way of commuting longer distances between towns and cities. Rising fuel and train ticket prices coupled with an increasing environmental awareness are all helping to boost its popularity.

Youth mindset for open sex:

Traditional Indian values about male female relationships, marriages have had a severe blow from the pornography available on internet. Online friendships sometimes turn into violent rapes on actual meeting. Youth is totally confused who want to be westernized but their deep rooted Indian values stops them to be free.

Flirting and sexting – sharing sexually explicit photos, videos, and chat by cell phone or internet – is fairly commonplace among young people. In general, 28% Indians flirted with someone other than partner.

Youth mindset for marriage:

Survey reveals that most of the Indians (80%) have used a mobile phone or internet for weddings; Along with that romances occur on the internet without physical or visual contact (Wiseman, 1998), which may end to marriage.

Media using youth as audience for advertising:

Mobile and internet are clearly becoming a mass medium in India. The advanced technology enables advertisers to discover their target audience on mobile Internet sites in a measurable manner. It is a well known fact that the Youth segment in India is one of the most lucrative consumer groups for marketers around the world.

Youth for politics:

As Dante Motes, a Bronx high school student who works at Youth Force, points out, "Instead of young people wanting to fight or just hanging out, they should be down and make a voice for themselves. If they don't, then who's going to be talking for us?" Such type of activity is gaining in India also.

Youth for voice rise:

Many youth argue that they don't need to be protected. Many believe that they should have access to all information, without it being censored by the obscenity and indecency laws theoretically meant to protect them. While allowing youth to access censored material like hard core pornography is highly debatable; it does suggest that perhaps we should start listening to youth, and their thoughts about the media.

4. Youth and health:

The ability to interact with various people without having to leave the comforts of their house leads to lack of social competencies and interpersonal skills to deal with the real world [8]. Internet addiction makes a person more aggressive, frustrated and isolated and leads to certain physical symptoms, such as back aches, dry eyes, migraine and other physical symptoms. Excessive use of mobiles can cause **Tumors And Acoustic Neuroma** - a study by a team from the Karolinska Institute found.

It also decreases the user's physical identity amongst their families, peers and co-workers. People tend to become isolated from the rest of the world as they continue to live in a fantasy of unreal identities. This is causing concern within society as people are thought to lose their sense of reality, interaction, and their personal identity. Online sharing has one of the most disadvantages of keeping all you have shared, in memory for ever. In reality, with passage of time, life changes and man gets matured and then all past memories simply disturb present life.

Virtual and multiple identities have further multiplied the human psychological problems as users lose touch with the sense of reality. This in turn can lead to depression about themselves, as once reality sets in that they are not who they created online, people can feel despondent about their lives and those in it.

On the positive side, young people may self-examine themselves by changing their character as more appealing. People can enhance their self esteem by forming on-line friendships and relationships.

5 Conclusion

We live today in an ever-increasingly hyper-interconnected world, a global society of communicative interactions and exchanges that stimulates profound cultural transformations and realignments. Indian youth is the most influenced segment that uses the innovative methods of communication and interaction. Young generation of India is very eager to set their career and make money sideways. With high traditional values in

Indian society, cultural diffusion is occurring. India is also going through contradicting cultural changes. Our older generation is not comfortable with the way people work, interact, relate, learn, process information and view their community and their standing in the world. Within business, the changes are getting more and more adaptable but with the Indian tradition living styles, the education system still needs time for adaptation.

6. References:

- [1] Richard Process Appelbaum, William J. Chambliss "Sociology"
- [2] Greetz, Clifford "The impact of culture on the Concept of man", 1968, The interpretation of culture, New York Basic books.
- [3] <http://vitalanalytics.in>
- [4] Indian leading newspaper "Times of India"
- [5] [TCS Generation Web 2.0 Trends 2008-09](#) Posted on August 04th, 2009

- [6] Bellis, M. (2005) "Inventors of the Modern Computer" retrieved September 3, 2005, from <http://inventors.about.com/library/weekly/aa031599.htm/>
- [7] Seel, J. (1997) "Plugged in, spaced out, and turned on: electronic entertainment and moral mindfields," Academic Search Elite vol.179, no.3, pp. 1-11 retrieved August 30, 2005, from Academic Search Elite database
- [8] <http://www.futureofchildren.org>
- [9] M/Cyclopedia of New Media (2004) "Virtual Identities" retrieved August 28, 2005, from http://newmediawiki.ci.qut.edu.au/index.php/Virtual_Identities/

Reliable and Energy Aware QoS Routing Protocol for Mobile Ad hoc Networks

V.Thilagavathe

Lecturer, Department of Master of Computer Applications,
Institute of Road & Transport Technology

K.Duraiswamy

Dean
K.S. Rangasamy College of Technology,
Tiruchengode.

Abstract—In mobile ad hoc networks (MANETs), always there is a tradeoff between reliability and energy consumption because in order to achieve maximum reliability, the maximum energy has to be consumed. But most of the existing works either concentrate on energy consumption or reliability, but very rarely both are taken into consideration. In this paper, we propose to develop a reliable and energy aware Quality of Service (QoS) Routing Protocol for MANETs to provide a combined solution for both energy consumption and reliability. In this protocol, multiple disjoint paths are determined for a source and destination and the routes are selected based on Route availability (RA) which is estimated from link availability (LA) and total energy consumed (TE) during the transmission of packets. By simulation results we show that the proposed protocol achieves better packet delivery ratio with reduced energy consumption and delay.

Keywords—Mobile ad hoc networks (MANETs); Quality of Service (QoS); Link availability (LA); ROUTE ERROR (RERR); ROUTE REQUEST (RREQ).

I. INTRODUCTION

A. Mobile Ad-Hoc Network (MANET)

A Mobile Ad-Hoc Network (MANET) is a self-configuring network of mobile nodes connected by wireless links, to form an arbitrary topology. The nodes are free to move arbitrarily. Thus, the network's wireless topology may be random and may change quickly. Such a network may operate in a standalone fashion, or may be linked to the larger Internet. An ad Hoc network is formed by sensor networks consisting of sensing, data processing, and communication components. Due to its deficiency in infrastructure support, each node acts as a router, forwarding data packets for other nodes [1]. Its application area includes Tactical Networks, Emergency Services, Commercial Environments Educational Applications and Entertainment [2].

B. Routing In MANET

Routing is the process of selecting paths in a network along which to send network traffic. Routing is performed for many kinds of networks, including the telephone network (Circuit switching), electronic data networks (such as the Internet), and transportation networks

- Nodes in traditional wired networks do not route packets, while in MANET every node is a router.
- Nodes transmit and receive their own packets and also forward packets for other nodes.

- Due to mobile nodes, topologies are dynamic in MANET, but are relatively static in traditional networks.
- Connectivity and interference are indicated by link layer information.
- A traditional router has an interface for each network to which it connects, while a MANET "router" has a single interface.
- Routed packet sent forward during transmission also gets transmitted to the previous transmitter.
- MANETs may have gateways to fixed network, but are normally "stub networks".

C. QoS Routing

QoS refers to the ability of a network to provide better service to selected network traffic over various technologies, including Frame Relay, Asynchronous Transfer Mode (ATM), Ethernet and 802.1 networks, SONET, and IP-routed networks that may use any or all of these basic technologies. Dedicated bandwidth, controlled jitter and latency (required by some real-time and interactive traffic), and improved loss characteristics are primary objective of QoS [4].

It is necessary to implement state-dependent, QoS-aware networking protocols, to enable QoS routing. A link weight expresses the available resources on a link. Though simple and reliable, flooding involves unnecessary communications and causes incompetent use of resources, particularly in the perspective of QoS routing that needs frequent distribution of multiple, dynamic parameters. Since all changes are not so important, monitoring changes via internet are not possible and desirable. Two possible changes are considered:

(1) Rare changes due to joining/leaving of nodes. In the current Internet, only this kind of topology changes is considered. Its dynamics are relatively well understood.

(2) Frequent changes, which are typically related to the consumption of resources or to the traffic flowing through the network [5].

1) Challenges of QoS Routing in Ad Hoc Networks:

- Dynamic varying network topology
- Imprecise state information

- Scarcity of resources
- Absence of communication infrastructure
- Lack of centralized control
- Power limitations
- Heterogeneous nodes and networks
- Error-prone shared radio channel
- Hidden terminal problem
- Insecure medium
- Other layers

2) Issues of QoS routing in MANET:

- ❖ To define accurately what QoS means in ad-hoc networks, is one of the basic issue. In wired networks, QoS usually means that the network guarantees something like a minimum bandwidth or maximum delay to a flow. In contrast, it is not possible to guarantee with certainty the bandwidth or delay to flows in an ad-hoc network when nodes are mobile and the links break and re-join in an unpredictable way. However, it is possible for the network to do the best it can to protect flows in predictable circumstances.
- ❖ Some Open Issues
 - QoS metric selection and cost function design
 - Multi-class traffic
 - Scheduling mechanism at source
 - Packet prioritization for control messages
 - QoS routing that allows preemption
 - Integration/coordination with MAC layer
 - Heterogeneous networks[8]
- ❖ QoS in ad-hoc network requires appropriate co-operation between various layers of the ad-hoc protocols, and it's difficult to achieve. Moreover, there is a wide range of applications and network conditions – implying that a single solution is doubtful to satisfy these varied requirements.
- ❖ Mobile ad hoc networks (MANET [1~2]) are characterized by high mobility and frequent link failures that ends in low throughput and high end-to-end delay. The function of a QoS routing strategy is to compute paths that are appropriate for different type of traffic generated by various applications while maximizing the utilizations of network resources. But the difficulty of finding multi-constrained paths has high computational complexity, and thus there is a need to use algorithms that resolve this difficulty [11].
- ❖ It is not only sufficient to find a route from a source to one or multiple destinations for QoS routing, also it has to satisfy one or more QoS constraints, typically, but not restricted to, bandwidth or delay [12].
- ❖ Homogeneous nodes in MANETs are assumed by existing routing protocols. i.e., all nodes have the same communication capabilities and characteristics. However, in many ad hoc networks, nodes are not the same. Some nodes have longer transmission range, larger transmission bandwidth, and are more reliable and robust than other nodes.
- ❖ Due to the nature constraints of MANETs, such as dynamic network topology, wireless communication link and limited process capability of nodes, it is difficult to find reasonable QoS routes in MANETs [13].
- ❖ Fast changing of topology and the lack of global information, are major challenges of a MANET. An MANET has only its adjacent information due to its transmission capability, such as transmission range, battery power, and so on. These result in complexity for developing an efficient routing protocol for MANETs [14].

D. Problem Identification and Proposed Solution

Always there is a tradeoff between reliability and energy consumptions because when reliability is maximum energy consumption also rises to maximum. Most of the existing works either concentrate on energy consumption or reliability, but both are not taken into consideration. For example, the error-aware feature of ECSRP [20] helps to reduce the energy consumption due to the retransmission of packets, but doesn't provide any solution for reliability. Prediction based link availability estimation [21] provides solution for reliability. It tries to predict the probability that an active link between two nodes will be continuously available for a predicted period, which is obtained based on the current node's movement, and this doesn't concentrate on energy consumption. Hence we have to design an algorithm in such a way that it need to provide a combined solution for both energy consumption and reliability.

In this paper, we develop a reliable and energy aware QoS Routing (REQR) Protocol for MANETs. The solution can be obtained by providing trade off between link availability and given threshold values.

In this protocol, the routes are selected based on Route availability (RA) which is estimated from link availability (LA) and total energy consumed (TE) during the transmission of packets. It will first select the routes satisfying the condition $RA > Th1$, where $Th1$ is the minimum threshold value for route stability. Among the selected routes, again it will choose the routes satisfying the condition $TE < Th2$, where $Th2$ is the minimum threshold value for total energy consumed.

If both the conditions are satisfied, then selected route is appropriate, and from this it's clear that selected route will have maximum link availability and minimum energy consumption. The Link availability is estimated based on the relative mobility of the nodes and the received signal strength.

II. RELATED WORKS

Chunxue Wu et al [11] have proposed a novel QoS Multipath Path Routing Mobile in MANET. Routing algorithm called Ad hoc on-demand Multipath routing was introduced, that provides quality of service (QoS) support, (Q-AOMDV) in terms of bandwidth, hop count and end-to-end delay in mobile ad hoc networks. The consequences validate that Q-AOMDV provides QoS support in mobile ad hoc wireless networks with high reliability and low overhead".

Mamoun Hussein Mamoun [16] have proposed, novel algorithm to routing called NRA, in mobile ad hoc networks, that allows the network layer to adjust its routing protocol dynamically based on SNR along the end-to-end routing path for each transmission link.

B. Sun et al [17] have proposed, Fuzzy Controller Based QoS Routing Algorithm with a Multiclass Scheme for MANETs. They presented a fuzzy controller based QoS routing algorithm with a multiclass scheme (FQRA) in mobile ad hoc networks. Comparison with "crisp" versions of the fuzzy algorithm to isolate the contributions of fuzzy logic, including applications of fuzzy control to power consumption and directional antennas in MANETs, are included in future works. They have also planned to compare FQRA with other QoS routing algorithm.

Fujian Qin et al [18] have proposed *Multipath Based QoS Routing* in MANET. A multipath source routing protocol with bandwidth and reliability guarantee is proposed. The protocol selects several multiple alternate paths which meet the QoS requirements, in route discovery phase, and to compromise between load balancing and network overhead, the ideal number of multipath routing is achieved. It can effectively deal with route failures similar to DSR, in route maintenance phase.

Md.Mamun-Or-Rashid et al [19] have proposed Link Stability and Lifetime Prediction Based QoS Aware Routing for MANETs. QoS aware routing problem is been formulated to maximize the link stability and lifetime whereas to minimize the cost. Their algorithm selects the best path in terms of link stability and lower cost lifetime prediction to minimize blocking probability along with QoS support. LSLP can reduce blocking probability up to 20% than that of other algorithms. LSLP decreases network life time a little than that of CLPR at the cost of better network performance and throughput as it will reduce packet loss due to network disconnection. Their proposed method formulates a tradeoff between link stability and cost which will guarantee a disruption free communication for transmission.

III. RELIABLE AND ENERGY AWARE QOS ROUTING PROTOCOL

A. Overview

Link availability, provides the basis for path selection based on the probability the path will remain available over a specified interval of time. It is a probabilistic model which predicts the future status of a wireless link. The reliability of a path depends on the availability of all links constituting this path [7]. A prediction based link availability estimation is used

which lets a node to first predict a continuous time period that a currently available link will last from initial time, by assuming that both nodes of the link will keep their current movements (i.e., speed and direction) unchanged [21].

To prolong the lifetime of MANETs, the energy consumption must be limited. In wireless channels, the channel condition also affects the power consumption [20].

Hence, in our protocol, we have taken both link availability and energy consumption as QoS parameters in route discovery.

The random ad-hoc mobility model is a continuous-time stochastic process that characterizes the movement of nodes in a two-dimensional space. Based on the mobility model, each node's movement consists of a sequence of random length intervals called *mobility epochs*, during which a node moves in a constant direction at a constant speed [7]. In our protocol, the following assumptions are made for the epoch length.

1. Mobility epoch lengths are exponentially distributed with mean λ^{-1} , i.e.

$$\bullet \quad E(x) = P\{RLV \leq x\} = 1 - e^{-\lambda x} \quad (1)$$

• Node mobility is uncorrelated.

Where, λ = epoch length,

RLV = Random length interval= λ

P = probability for the two nodes to move close to each other after changing their movements.

B. Estimation of QoS Metrics

Based on above assumptions, the availability of link is defined as,

$$L_A(P_r) = \{A_{t0}, t_0 + P_r\} \quad (2)$$

it indicates the probability that the link will be continuously available from time t_0 to $t_0 + P_r$.

Where, P_r = continuous time period.

So, the expression for link availability is derived as:

$$L_{\min}(P_r) = \frac{1 - e^{-2\lambda Pr}}{2\lambda Pr} + \frac{\lambda Pr e^{-2\lambda Pr}}{2} \quad (3)$$

Where, $e^{-2\lambda Pr}$ is the probability that nodes keep their movements same.

The basic assumption of Energy Efficient Routing Protocol in MANETs includes transmission energy model using Shadowing.

• Transmission energy per scale can be expressed as,

$$P_t = \lambda * d^r \quad (4)$$

Where, λ is determined by frequency of radio, receiver threshold and signal-to-noise threshold.

- Transmitting RTS and CTS consumes the following total energy,

$$RTS_E = \lambda * d_{\max} * T_{RTS} \quad (5)$$

$$CTS_E = \lambda * d_{\max} * T_{CTS} \quad (6)$$

- Data and Acknowledgement is transmitted using the following energy,

$$DATA_E = \lambda * d_{ij} * T_{DATA} \quad (7)$$

$$ACK_E = \lambda * d_{ij} * T_{ACK} \quad (8)$$

- d_{\max} is maximum transmission range
- d_{ij} is distance between node i and node j
- T_{RTS} , T_{CTS} , T_{DATA} , T_{ACK} are transmission time of RTS, CTS, DATA, and ACK packets.

- Then transmitting a packet along a path requires the total energy,

$$E_{SD} = \sum_{t=1}^{m-1} \gamma * d_t(t+1) * (T_{DATA} + T_{ACK}) + \sum_{t=1}^{m-1} \gamma * d_{\max} * (T_{RTS} + T_{CTS}) + \sum_{t=1}^{m-1} P_r * (T_{DATA} + T_{ACK} + T_{RTS} + T_{CTS}) \quad (9)$$

P_r = receiving energy per second and is assumed to be regardless of the packet type.

C. Route Discovery

AOMDV is an on-demand routing protocol that builds multiple routes using request/reply cycles. When the source needs a route to the destination but no route information is known, it floods the ROUTE REQUEST (RREQ) message to the entire network. Because this packet is flooded, several duplicates that traversed through different routes reach the destination. The destination node selects multiple disjoint routes and sends ROUTE REPLY (RREP) packets back to the source via the chosen routes. The purpose of computing alternate paths for a source node is that when the primary path breaks due to node movement, one of the alternate paths can then be chosen as the next primary path and data transmission can continue without initiating another route discovery [11].

In our proposed protocol, RREQ message additionally includes hop count, link availability as well as energy consumed so as to select the primary path in all the available paths while message is broadcasted upon receiving a route request to the destination. Similarly RREP message also contains the metrics talked above. The mobile ad hoc network is modeled as a graph $G = (N, L)$, where N is a finite set of nodes and L is a set of bi-directional links. The protocol will only use bi-directional links, so any unidirectional links are

omitted. The QoS metrics are evaluated by energy (E) and path availability (L).

In multi-path routing protocol, the hop count is also considered as a metric to evaluate the QoS. We conclude the following expressions to calculate the energy, link availability and hop count.

$$P_{\min}(\text{Pr}) = \sum L_{\min}(\text{Pr}) \quad (10)$$

$$E_{SD} = \sum_{t=1}^{m-1} \gamma * d_t(t+1) * (T_{DATA} + T_{ACK}) + \sum_{t=1}^{m-1} \gamma * d_{\max} * (T_{RTS} + T_{CTS}) + \sum_{t=1}^{m-1} P_r * (T_{DATA} + T_{ACK} + T_{RTS} + T_{CTS})$$

$HC(\text{path}(m, n))$ = Hop count for the path $m - n$

Using the above expressions we can obtain the metric values of each path in the network and find an evaluation method for path preference probability P_p , which aims at finding a path that satisfies certain requirements such as energy, link availability, and hop count.

$$P_p = \frac{P_{\min}}{TE * HC} \quad (11)$$

Where, P_p = path preference probability, TE = Total energy, HC = Hop count, P_{\min} = path availability,

By calculating and comparing the P_p in the available paths, the path which has higher path preference probability will be selected for data transmission.

D. Route Maintenance

In our routing protocol, when a node fails to deliver the data packet to the next hop of the route, it considers the link to be disconnected and sends a ROUTE ERROR (RERR) packet to the upstream direction of the route. The route to the source and the immediate upstream and downstream nodes of the broken link is contained in RERR message. The source removes every entry in its route table that uses the broken link (regardless of the destination), upon receiving the RERR packet. The source uses the remaining valid route to deliver data packets, if one of the routes of the session is removed. Some or all of the routes may break due to node mobility and/or link and node failures, after a source begins sending data along multiple routes [11].

During data transmission through the primary path, whenever the link availability of one or more links becomes less than a minimum threshold value,

$$\text{i.e., } P_{\min} < Th_{\min}$$

Then, ROUTE ERROR (RERR) packet is sent to source node along the route.

From the multiple disjoint paths determined, source node will fetch the next better path and re-route the traffic through

this path. Normally, route error packet is sent when there is route break, and the recovery is carried out after that. Here, the recovery is performed proactively before the route break or route failure.

IV. SIMULATIONS AND RESULTS

A. Simulation Model and Parameters

We use NS2 [22] to simulate our proposed protocol in our simulation; the channel capacity of mobile hosts is set to the same value: 2 Mbps. We use the distributed coordination function (DCF) of IEEE 802.11 for wireless LANs as the MAC layer protocol. It has the functionality to notify the network layer about link breakage.

In our simulation, mobile nodes move in a 1000 meter x 1000 meter region for 100 seconds simulation time. We assume each node moves independently with the same average speed. All nodes have the same transmission range of 250 meters. The network size is varied as 25, 50, 75 and 100 nodes and the pause time of the mobile node is varied as 0,10,20,30 and 40. The speed of the mobile node is set as 5m/s. The simulated traffic is Constant Bit Rate (CBR).

Our simulation settings and parameters are summarized in table I

TABLE I. SIMULATION SETTINGS

No. of Nodes	50
Area Size	1000 X 1000
Mac	802.11
Radio Range	250m
Simulation Time	100 sec
Traffic Source	CBR
Packet Size	512
Mobility Model	Random Way Point
Speed	5m/s
Pause time	0,10,20,30 and 40
Transmit Power	0.360 w
Receiving Power	0.395 w
Idle Power	0.335 w
Initial Energy	5.1 J

B. Performance Metrics

We compare our REQR protocol with the LSLP [19] protocol.

We evaluate mainly the performance according to the following metrics.

Control overhead: The control overhead is defined as the total number of routing control packets normalized by the total number of received data packets.

Average end-to-end delay: The end-to-end-delay is averaged over all surviving data packets from the sources to the destinations.

Average Packet Delivery Ratio: It is the ratio of the number of packets received successfully and the total number of packets sent.

Drop: It is the number of packets dropped.

Average Energy: It is the average energy consumption of all nodes in sending, receiving and forward operations.

C. Results

A. Based on Pause time

In our initial experiment, we vary the pause time as 0,10,20,30 and 40.

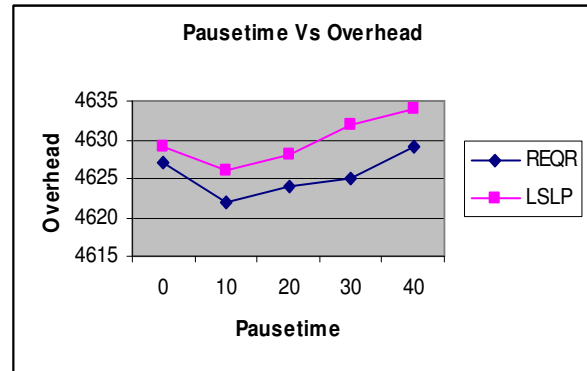


Figure 1. Pausetime Vs Overhead

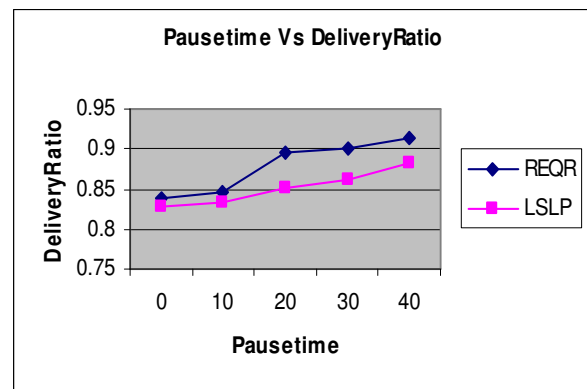


Figure 2. Pausetime Vs Packet Delivery Ratio

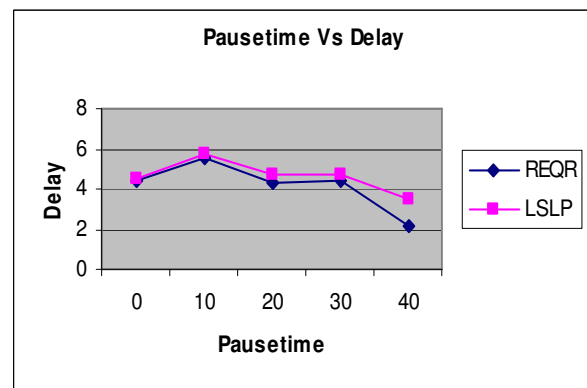


Figure 3. Pausetime Vs Delay

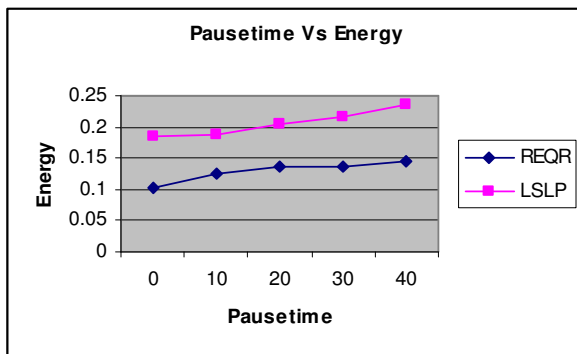


Figure 4. Pausetime Vs Energy

From Figures 1, we can ensure that the control overhead is less for REQR when compared to LSLP.

Figure 2 presents the packet delivery ratio of all the protocols. Since the packet drop is less and the throughput is more, REQR achieves good delivery ratio, compared to LSLP

From Figure 3, we can see that the average end-to-end delay of the proposed REQR protocol is less when compared to the LSLP protocol.

Figure 4 shows the results of energy consumption for the pause time 0,10,20...40. From the results, we can see that REQR scheme has less energy than the LSLP, since it has the energy efficient routing.

B. Based On Number of Nodes

In the second experiment, we vary the number of nodes as 25, 50, 75 and 100.

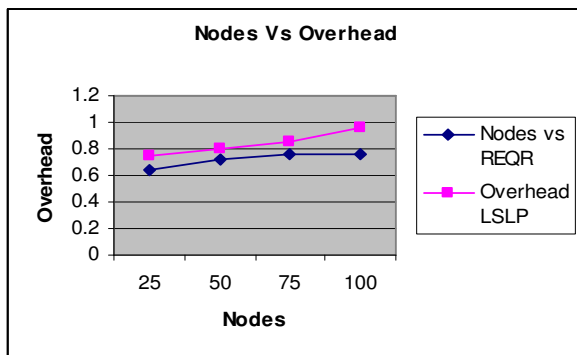


Figure 5. Nodes Vs Overhead

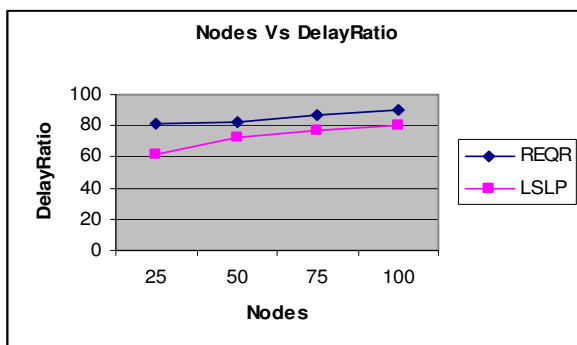


Figure 6. Nodes Vs Packet Delivery Ratio

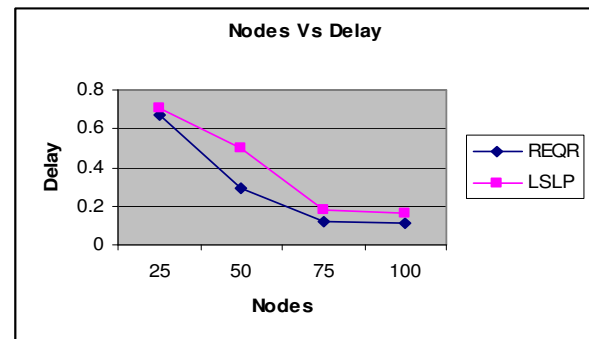


Figure 7. Nodes Vs Delay

From Figure 5, we can ensure that the control overhead is less for REQR when compared to LSLP. Figure 6 presents the packet delivery ratio of all the protocols. From the figure, we can observe that REQR achieves good delivery ratio, compared to LSLP. From Figure 7, we can see that the average end-to-end delay of the proposed REQR protocol is less when compared to the LSLP protocol.

V. CONCLUSION

In this paper, we have developed a reliable and energy aware Quality of Service (QoS) Routing Protocol for MANETs to provide a combined solution for both energy consumption and reliability. In this protocol, the routes are selected based on Route availability (RA) and total energy consumed (TE) during the transmission of packets. Link availability, provides the basis for path selection based on the probability that the path will remain available over a specified interval of time. Initially multiple disjoint paths are determined for a source and destination. Using these metrics, we obtain the combined metric value of each path in the network and find an evaluation method path preference probability P_p , which aims at finding a path that satisfies the requirements such as energy, link availability, and hop count. Then the path which has higher path preference probability will be selected as a primary path for data transmission. During data transmission through the primary path, whenever the link availability of one or more links becomes less than a minimum threshold value, the ROUTE ERROR packet is sent to source node along the route. From the multiple disjoint paths determined, source node will fetch the next better path and re-route the traffic through this path. So the recovery is performed proactively before the route break or route failure. By simulation results we have shown that the proposed protocol achieves better packet delivery ratio with reduced energy consumption and delay.

REFERENCES

- [1] T. Hara, "Effective replica allocation in ad hoc networks for improving data accessibility", Proc. IEEE Infocom 2001, pp.1568-1576, 2001.
- [2] www.wikipedia.org
- [3] <http://datatracker.ietf.org/wg/manet/charter/>
- [4] Quality of Service (QoS) Networking, chapter 46, Internetworking Technology Overview, June 1999
- [5] X. Masip-Bruina,, M. Yannuzzib, J. Domingo-Pascuala, A. Fonteb, M. Curadob, E. Monteirob, F. Kuipersc, P. Van Mayhem, S. Avalloned, G.

- Ventred, P. Aranda-Gutierrez, M. Hollickf, R. Steinmetzf, L. Iannoneg, K. Salamatiang, "Research challenges in QoS routing", 2005.
- [6] <http://www.cs.ucr.edu/~csyiazti/cs260.html>
- [7] A. Bruce McDonald and Taieb Znati, "A Path Availability Model for Wireless Ad-Hoc Networks", May 1999.
- [8] <http://www.slideshare.net/Annie05/qo-s-routing-in-ad-hoc-networks-presentation>
- [9] YU-CHEE TSENG, WEN-HUA LIA and SHIH-LIN WU, "Mobile Ad Hoc Networks and Routing Protocols" Copyright © 2002 John Wiley & Sons, Inc., ISBNs: 0-471-41902-8 (Paper); 0-471-22456-1 (Electronic) .
- [10] Jean Walrand, "Implementation of QoS Routing for MANETs", June 30, 2007
- [11] Chunxue Wu, Fengna Zhang, Hongming Yang, "A Novel QoS Multipath Path Routing in MANET", 2010
- [12] Philipp Becker, "QoS Routing Protocols for Mobile Ad-hoc Networks", 2007.
- [13] Xiaojiang Du, "Delay Sensitive QoS Routing For Mobile Ad Hoc Networks", 2003.
- [14] Jitendranath Mungara, "Design and a New Method of Quality of Service n Mobile Ad Hoc Network", European Journal of Scientific Research, ISSN 1450-216X Vol.34 No.1 (2009), pp.141-149, © EuroJournals Publishing, Inc. 2009.
- [15] Kuei-Ping Shih, Chih-Yung Chang, Yen-Da Chen and Tsung-Han Chuang, "Dynamic bandwidth allocation for QoS routing on TDMA-based mobile ad hoc networks", 22 November 2005.
- [16] Mamoun Hussein Mamoun, "A Novel Routing Algorithm for MANET", International Journal of Electrical & Computer Sciences IJECS-IJENS Vol:10 No: 02, 2009.
- [17] B. Sun, C. GUI, Q. Zhang, H. Chen, "Fuzzy Controller Based QoS Routing Algorithm with a Multiclass Scheme For MANET", Int. J. of Computers, Communications & Control, ISSN 1841-9836, E-ISSN 1841-9844 Vol. IV (2009), No. 4, pp. 427-438.
- [18] Fujian Qin and Youyuan Liu, "Multipath Based QoS Routing in MANET", JOURNAL OF NETWORKS, VOL. 4, NO. 8, October 2009.
- [19] Md. Mamun-Or-RashidO and Choong Seon Hong, "LSLP: Link Stability and Lifetime Prediction Based QoS Aware Routing for MANET", 2006.
- [20] Liansheng Tan, Peng Yang, Sammy Chan, "An Error-aware and Energy Efficient Routing Protocol in MANETs", Proceedings of 16th International Conference On Computer Communications and Networks, 2007. ICCCN 2007.
- [21] Shengming Jiang, Dajiang He and Jianqiang Rao, "A Prediction-based Link Availability Estimation for Mobile Ad hoc Networks", Proceedings.Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM , 2001.

V. Thilagavathe received the B.E. degree in Computer Science and Engineering from the Bharathiar University in 1995, and the M.E. degree in Computer Science and Engineering from the Anna University, Chennai in 2004. Her research activity includes QoS routing and mobile ad hoc networks. From 1997 to 2000 she was working in K.S.Rangasamy College of Technology, Tiruchengode. From 2002 to 2005 she was in K.S.R. College of Engineering, Tiruchengode. She is currently working in Institute of Road and Transport Technology, Erode as Lecturer of Master of Computer Applications Department since 2006. She is a life member of ISTE and CSI.

Dr. K.Duraiswamy (SM) received his B.E. degree in Electrical and Electronics Engineering from P.S.G. College of Technology, Coimbatore, Tamil Nadu in 1965 and M.Sc.(Engg) degree from P.S.G. College of Technology, Coimbatore, Tamil Nadu in 1968 and Ph.D. from Anna University, Chennai in 1986. From 1965 to 1966 he was in Electricity Board. From 1968 to 1970 he was working in ACCET, Karaikudi, India. From 1970 to 1983, he was working in Government College of Engineering, Salem. From 1983 to 1995, he was with Government College of Technology, Coimbatore as Professor. From 1995 to 2005 he was working as Principal at K.S. Rangasamy College of Technology, Tiruchengode and presently he is serving as Dean in the same institution. He is interested in Digital Image Processing, Computer Architecture and Compiler Design. He received 7 years Long Service Gold Medal for NCC. He is a life member in ISTE, Senior member in IEEE and a member of CSI.

A DYNAMIC APPROACH TO DEFEND AGAINST ANONYMOUS DDOS FLOODING ATTACKS

Mrs. R. ANUREKHA

Lecturer, Dept. of IT
Institute of Road and Transport Technology,
Erode, Tamilnadu, India.

Dr. K. DURAISWAMY

Dean, Department of CSE
K.S.Rangasamy College of Technology,
Tiruchengode, Namakkal, Tamilnadu, India.

A.VISWANATHAN

Lecturer, Department of CSE
K.S.R.College of Engineering,
Tiruchengode, Namakkal, Tamilnadu, India

Dr. V. P. ARUNACHALAM

Principal, SNS College of Technology,
Coimbatore, Tamilnadu, India

A. RAJIV KANNAN

Asst.Prof, Department of CSE
K.S.R.College of Engineering,
Tiruchengode, Namakkal, Tamilnadu, India.

K. GANESH KUMAR

Lecturer, Department of IT
K.S.R.College of Engineering,
Tiruchengode, Namakkal, Tamilnadu, India

Abstract: Several IP traceback schemes have been proposed to trace DoS/DDoS attacks that abuse the internet. A mechanism for IP traceback based on the geographic information rather than the traditional IP address information was proposed in [1], for 8 directions in a planar environment. Extension of this two dimensional directed geographical traceback to $2n$ [$n \geq 4$] directions is also available [2].

In this paper, the DGT scheme has been generalized to three dimensions, with all routers in a spherical environment in tune with reality. A traceback algorithm, called Direction Ratio Algorithm (DRA) enables IP traceback with robustness and fast convergence.

Keywords: IP traceback, spherical environment, DRS (Direction Ratio Set), DRA (Direction Ratio Algorithm).

1. INTRODUCTION

DDoS attacks continue to plague the internet, due to the availability of a plethora of attacking tools (TFN, Trin00 and stacheldraht) [3]. Since DDoS attacks rely on anonymity, it follows that a solution must eliminate some of the anonymity of the hosts. Finding the source of the spoofed packets, called the IP traceback problem is one of the hardest security problems needing redressal.

Among several traceback schemes, the directed geographical traceback (DGT) is based on geographical

information, rather than the traditional IP address [1]. This scheme has been extended to 2^n ($n \geq 4$), directions in a planar environment [2], where all the routers and devices are assumed to be coplanar, which is not always true.

In this paper, we have generalized DGT to three dimensions, where the true spherical topology of the geographical globe is taken into consideration for the traceback.

All the advantages (like robustness, fast convergence, independence etc.) of the two dimensional DGT are available in the three dimensional scheme as well. The basic assumptions about the traffic and the network are the same as in [1].

The rest of this paper is organized as follows. In section II, the spherical topology of the routers is introduced in normalized coordinates. Concept of DRS (Direction Ratio Set) & the uniqueness theorem are discussed in sections III & IV. Several options of NDRS (Neighborhood Direction Ratio set) and DRA (Direction Ratio Algorithm) traceback are described in sections V & VI. Limitations are discussed in section VII, while in section VIII conclusions and future prospects are detailed.

2. GEOGRAPHICAL TOPOLOGY OF THE EARTH:

Referred to rectangular axes, OX, OY, OZ, the earth can be, geographically considered as a sphere, having the equation,

$$X^2 + Y^2 + Z^2 = a^2 \quad (2.1)$$

With points A, B, C having coordinates (a,0,0), (0,a,0) and (0, 0, a) respectively

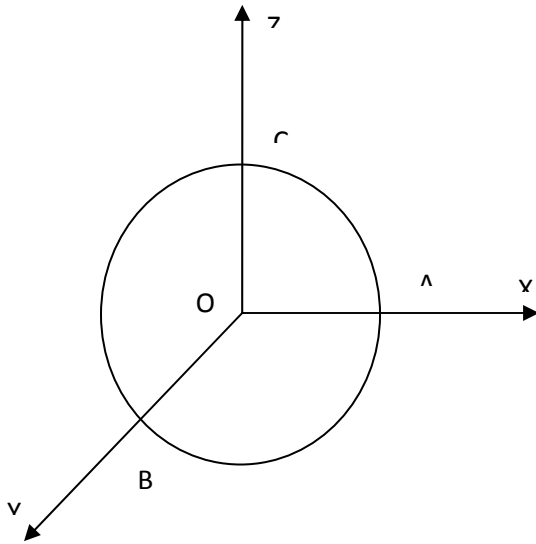


FIGURE 2.1-TOPOLOGY OF EARTH

Origin is at the centre & 'a' is the radius of the earth. Making the transformation

$$X=ax, Y=ay, Z=az \quad (2.2)$$

Eq. (2.1) gives

$$x^2 + y^2 + z^2 = 1 \quad (2.3)$$

where the metric unit is the radius of the earth.

Alternatively, assuming the ellipsoidal topology of the earth in the form

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = 1 \quad (2.4)$$

where under the transformation.

$$X=ax, Y=by, Z=cz \quad (2.5)$$

Eq. (2.4) gives

$$x^2 + y^2 + z^2 = 1 \quad (2.3)$$

Hence in our traceback study, the routers R_i are at chosen points

$P(x_i, y_i, z_i)$ on Eq. (2.3) where

$$x_i^2 + y_i^2 + z_i^2 = 1 \quad \text{for all } i.$$

3. CONCEPT OF DIRECTION RATIO SET (DRS) AT A ROUTER POINT.

The direction of a line in space, is indicated by their direction cosines ($\cos\alpha, \cos\beta, \cos\gamma$) where α, β, γ are the angles which the line makes with positive directions of the axes. (Refer Fig 3.1).

We can show

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1 \quad (3.1)$$

for all direction cosines (d.c).

The d.c being cumbersome fractions / irrationals in $[-1, 1]$, are not suited for IP traceback.

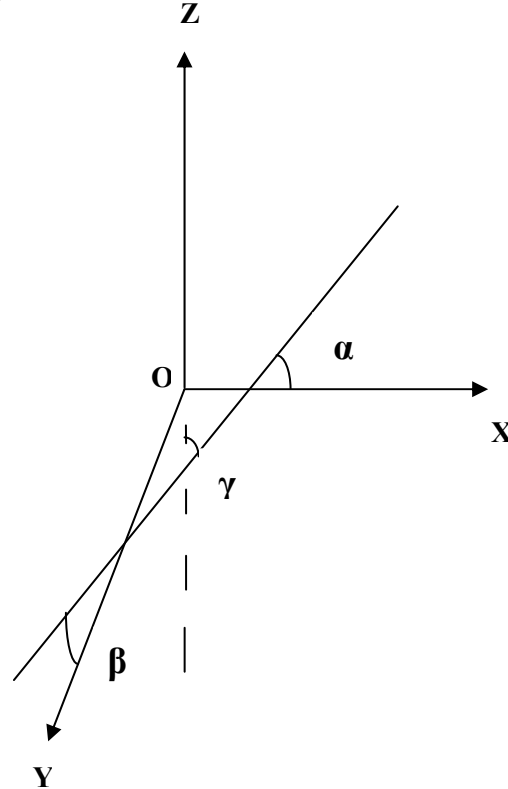


FIGURE 3.1 – DIRECTION ANGLES OF A LINE IN SPACE

Hence, we use proportional quantities to d.c, called direction ratios (d.r), denoted by (a, b, c) where a, b, c are integers with

$$\gcd(a, b, c) = 1 \quad (3.2)$$

Direction Ratio Set (DRS) at a router point R_0 , is the set D_0 of direction ratios

$$D_i = \{(a_i, b_i, c_i), i = 1 \text{ to } n\} \quad (3.3)$$

of its immediate neighbors R_i to R_n from R_0 (Refer fig 3.2).

Note that all router points R_i for $i = 0$ to n all lie on the unit sphere.

In contrast to two dimensional DGT, we can prove that, for any specific direction ratio (a_i, b_i, c_i) at R_0 , there is a unique router R_i on the sphere.

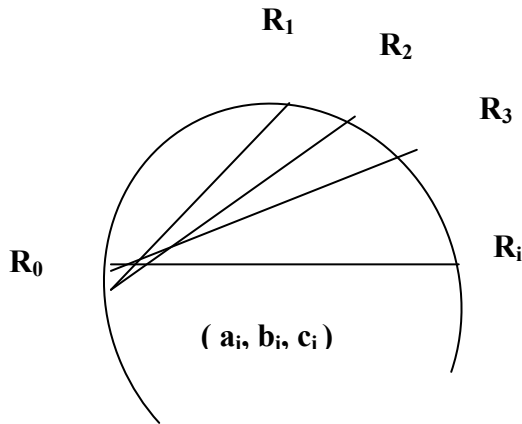


FIGURE 3.2 – DR SET FROM ROUTER R_0

4. UNIQUENESS THEOREM

A. Statement:

If (x_0, y_0, z_0) are the coordinates of router R_0 , then there is a unique router $R_i(x_i, y_i, z_i)$ in the directions $R_0 R_i$, with d.r (a_i, b_i, c_i) where

$$x_i = x_0 + a_i r, \quad y_i = y_0 + b_i r, \quad z_i = z_0 + c_i r \quad (4.1)$$

with

$$r = -2(a_i x_0 + b_i y_0 + c_i z_0) / (a_i^2 + b_i^2 + c_i^2) \quad (4.2)$$

B. Proof:

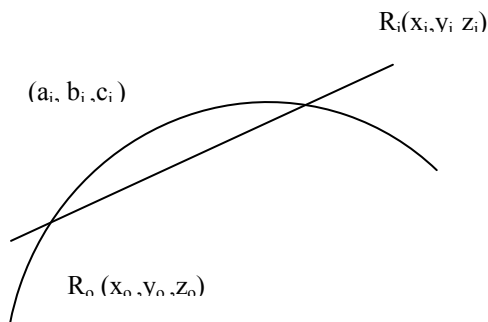


FIGURE 4.1 – (1-1) CORRESPONDENCE OF (a_i, b_i, c_i) AND R_i

The point R_i in parametric form is

$$x_i = x_0 + a_i r, \quad y_i = y_0 + b_i r, \quad z_i = z_0 + c_i r \text{ and lies on} \quad (4.3)$$

$$\therefore x_i^2 + y_i^2 + z_i^2 = 1 \quad (4.4)$$

Substituting in Eq. (4.4) and simplifying we get

$$r = -2(a_i x_0 + b_i y_0 + c_i z_0) / (a_i^2 + b_i^2 + c_i^2) \quad (4.5)$$

Thus there is a (1-1) correspondence between

$D_i = (a_i, b_i, c_i)$ (the d.r.) and the points

$R_i = (x_i, y_i, z_i)$ on the sphere

except when

$$a_i x_0 + b_i y_0 + c_i z_0 = 0 \quad (4.6)$$

when the direction is that of the tangent line at R_0 .

This uniqueness makes the three dimensional IP traceback, a robust one, converging on a single packet.

5. NEIGHBORHOOD DIRECTION RATIO SET (NDRS) AT A ROUTER POINT.

In space, from any router point R_0 , there are infinite directions, all of which, by uniqueness theorem give distinct, infinitely many, possible router points R_i on the unit sphere.

It is needless/ impossible for routers to know the d.r of all its successors. To reduce the router overhead, we introduce the concept of NDRS (Neighborhood Direction Ratio Set) which alone it should know.

In general, the direction ratio triad of integers (a_i, b_i, c_i) are allowed to take values given by

$$0 \leq a_i / b_i / c_i \leq n, \quad n \in \mathbb{N} \quad (5.1)$$

then $d(n)$, number of directions from R_0 satisfies the inequality

$$(2n)^3 < d(n) < (2n+1)^3 \quad (5.2)$$

due to the weeding out of redundant direction ratios from the total set.

The choice of n , and hence $d(n)$, depends on the field width reserved for each d.r triad in the packet header. It is easily verified that for a field width allotment of $3(m+1)$ bits for a d.r triad, the range is

$$0 \leq a_i / b_i / c_i \leq n \quad (5.3)$$

where $n = 2^m - 1$ and

$$(2n)^3 < d(n) < (2n+1)^3 \quad (5.4)$$

Specifically, for a field of 6 bits for a d.r triad (including 3 sign bits),

$$0 \leq a_i / b_i / c_i \leq 1 \text{ and } 8 < d(1) < 27$$

We can show that $d(1) = 13$ and the 13 d. r. are in Table 5.1 shown below.

6. THREE DIMENSIONAL TRACEBACK PROCEDURES.

Assuming that for every router the NDRS has been uniformly chosen, So that a Uniform field width is needed for the d.r marking, the traceback procedure is as follows: (for 13 directions, we need 6 bits/d.r).

Let $D_i = (a_i, b_i, c_i)$ be the D.R triad at router R of direction RR_i . Then the Direction Ratio Algorithm (DRA) is as follows:

TABLE 5.1 DIRECTION RATIOS OF D (1)

D_i	Directional Ratios	D_7	$(-1,0,1)$
D_1	$(1,0,0)$	D_8	$(0,1,1)$
D_2	$(0,1,0)$	D_9	$(0,-1,1)$
D_3	$(0,0,1)$	D_{10}	$(1,1,1)$
D_4	$(1,1,0)$	D_{11}	$(-1,1,1)$
D_5	$(-1,1,0)$	D_{12}	$(1,-1,1)$
D_6	$(1,0,1)$	D_{13}	$(1,1,-1)$

A. Marking procedure at router R

For each packet w , append D_i to w .

B. Path reconstruction at victim V

For any packet w from attacker, extract D.R list (D_1, D_2, \dots) from the suffix of w .

Unique traceback is now possible using the results (4.1) and (4.2).

If $(D_{n-1}, D_{n-2}, \dots, D_0)$ are the n suffixes of w during the n hops from R_n to R_0 then the path is constructed as in fig 6.1

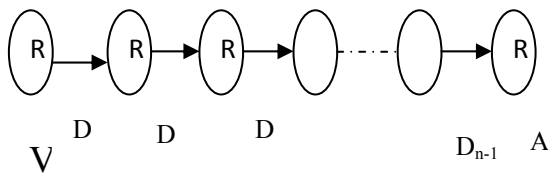


FIGURE 6.1 – TRACEBACK CONSTRUCTION

7. PERFORMANCE COMPARISON

A. COMPARISON OF DGT 16 WITH DGT 8

DGT 16 and DGT 8 being like schemes (the former, removing the directional constraints of the latter) they have equivalent advantages with respect to computational burden, scalability and mitigation capability of the attack, except for the fact that 16 directions are available now, with nil or negligible additional computations.

B. QUALITATIVE COMPARISON OF DGT 16 WITH OTHER TRACEBACK SCHEMES

Due to the totally different nature of DGT and other well known traceback schemes, involving packet marking or packet logging techniques, quantitative comparison of the various schemes is not possible. Hence in this section, we first present a qualitative comparison between DGT and other well known traceback schemes.

Success of any traceback scheme is determined by four key factors – computational overhead involved for packet marking, memory requirement for packet logging, scalability of the proposed scheme and the need for cooperation between other do mains.

The overhead of the DGT presented here is very light; The DGT scheme is also scalable. No Cooperation between different ISPs is required. Furthermore unlike PPM and SPIE, the scheme can be used to mitigate the effect of the attack while the attack is ragging on.

The comparison summary is in Table 7.1. The result as reported in table proves the superiority of directed geographical IP Traceback with respect to computational load, scalability and mitigation capability parameters over all other previously proposed schemes.

C. LIMITATIONS OF DRA

DRA is both robust and extremely quick to converge (on a single packet) and is independent. For 13 directions/router, the field /d.r is as small as 6 bits per hop. Yet there are limitations.

Apart from the router overhead incurred by appending data to packets in flight, since the length of the path is not known apriori, it is impossible to ensure that there is sufficient unused space in the packet for the complete list of d.r of the path.

This problem can be addressed by d.r sampling by the routers on the path, one at a time, instead of recording the entire path list of d.r.

8. CONCLUSION

We have generalized the ideal, two dimensional DGT, to real three dimensional DGT on a unit sphere. Concepts of DR, DRS and NDRS along with the uniqueness theorem have been introduced.

The DRA traceback is qualitatively robust, with fast convergence and independence. The storage issue is addressed through the DRSA traceback, (d.r sampling algorithm) which will be reported with further work, so as to make 3 dimensional, multidirectional, Geographical traceback more useful.

TABLE 7.1 PERFORMANCE COMPARISON OF VARIOUS TRACEBACK SCHEMES

Scheme		Author	Memory Requirements at Routers	Computational Burden	Scalability	Traceback	
						Time required	Number of packets required
Logging	SPIE	Strayer, et.al	High	High	Poor	Low	Traced each packet
	Distributed-Log-based Scheme for IP Traceback	Jing , et.al	Fair	High	Good	Low	Low
PPM	ID-Based PPM for IPT	Tseng, et.al	NIL	High	Good	Fair	High
	ERPPM	LIU , et.al	NIL	High	Good	Medium	Low
DPM	Flexible Deterministic Packet Marking (FDPM)	Xiang , et.al	NIL	Medium	Good	Fair	Low
	Pi: A Path Identification Mechanism	Yaar, et.al	NIL	Light	Good	Fair	Fair
Other Approaches	A Real-Time Traceback Scheme - DDoS Attacks	Huang , et.al	NIL	Medium	Good	Fair	Fair
	Marking and Logging for Marking and Logging forIPT	Al-Duwairi, et.al	Medium	Medium	Good	Fair	Fair
DGT	2D 16 directional DGT	Kannan, et.al	NIL	Light	Good	Negligible	Traced each packet
	3D multi-directional DGT	Kannan, et.al	NIL	Light	Good	Negligible	Traced each packet

9. REFERENCES

- [1]. Zhiqiang Gao and Nirwan Ansari (2005), "Directed Geographical Traceback", IEEE transactions , IEEE paper 221-224.
- [2]. A.Rajiv Kannan, Dr.K.Duraiswamy (2008),"16 directional DGT with generalization to 2n (n>4) directions", IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.11.
- [3]. CERT, "Computer emergency response team, CERT advisory ca-2000-2001: Denial-of-service development http://www.cert.org/advisories/CA-2000-01.html,2000.
- [4]. S.Savage, D.Wetherall (2001), "Practical network support for IP Traceback" IEEE/ACM transactions on Networking Vol 9-pp 226-237.
- [5]. V.Padmanahan. (2001), "Determining the geographic location of internet hosts," ACM SIGMETRICS '01; Cambridge, MA., pp 324-325.
- [6]. V.Padmnabham. (2001), "An investigation of geographic mapping techniques for internet hosts", ACM SIGCOMM'01.,San Diego., CA., pp 173-185.
- [7]. P.Ferguson (1998), "Network ingress filtering: defeating DOS attacks which employ IP source address spoofing", RFC 2267.
- [8]. R.Govindan (2000), "Heuristics for internet map discovery", Proceedings of IEEE INFOCOM conference.,Tel Aviv.,Israel.
- [9]. B.Al-Duwairi and T.E.Daniela (2004), "Topology based marking", IEEE int. conf on computer comm. and networks.,(ICCCN).
- [10]. T.Baba and S.Matsuda (2002), "Tracing network attacks to their sources ", Proe .conf IEEE internet computing., Vol 6., No:2,pp 20-26.

AUTHORS PROFILE:



ANUREKHA R received B.E. and M.E degrees, from Madras University and Anna University in 1998 and 2004 respectively . She is currently working as a Lecturer in the Department of Information Technology at Institute of Road and Transport Technology, affiliated to Anna University. Her research interest includes Network and Security. She is also a member of ISTE.



DR.K.DURAI SWAMY received the B.E., M.Sc. and Ph.D. degrees, from the University of Madras and Anna Univ. in 1965,1968 and 1987 respectively. After working as a Lecturer (from 1968) in the Dept. of Electrical Engineering in Government College of Engineering, Salem - affiliated to Anna Univ. and as an Asst. professor (from 1983) in Government College of Technology, Coimbatore(Anna Univ.),and as a Professor and Principal (from 1995) at K.S.Rangasamy College of Technology (Anna Univ.). He has been working as a Dean in the Dept. of Computer Science and Engineering at K.S.Rangasamy College of Technology, Anna University since 2005. His research interest includes Mobile Computing, Soft Computing, Computer Architecture and Data Mining. He is a Sr. member of ISTE, SIEEE, CSI.



A.VISWANATHAN received the B.E., degree from the Anna University, Chennai and M.E degree from Anna University, Coimbatore. He is doing his research in Network Security. His area of interest includes Operating Systems and Object Analysis and Design. He is a student member of ISTE.



A. RAJIVKANNAN received the B.E. and M.E degrees, from Periyar Univ. and Anna Univ. in 2002 and 2004, respectively. After working as a Lecturer (from 2004) and he has been a Senior lecturer in the Dept. Of Computer Science and Engineering at K.S.R. College of Engineering affiliated to Anna Univ. since June 2008. His research interest includes Network and its Security especially in IP traceback & DDoS.

Other areas include Operating Systems and MANET. He is a member of ISTE. One of his research paper was published in International Journal of Computer Science and Network Security in November 2008.



K. GANESH KUMAR received the B.Tech., degree from the Anna University, Chennai. His Research Area includes Computer Networks and security in 2006 and M.E degree from Anna University, Coimbatore., Operating Systems and Object Analysis and Design. He is a student member of ISTE.

IJCSIS REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway
Assoc. Prof. N. Jaisankar, VIT University, Vellore, Tamilnadu, India
Dr. Amogh Kavimandan, The Mathworks Inc., USA
Dr. Ramasamy Mariappan, Vinayaka Missions University, India
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania
Dr. Junjie Peng, Shanghai University, P. R. China
Dr. Ilhem LENGILIZ, HANA Group - CRISTAL Laboratory, Tunisia
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India
Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India
Mrs Li Fang, Nanyang Technological University, Singapore
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India
Mr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand
Mr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India
Mr. Hayder N. Jasem, University Putra Malaysia, Malaysia
Mr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India
Mr. R. S. Karthik, C. M. S. College of Science and Commerce, India
Mr. P. Vasant, University Technology Petronas, Malaysia
Mr. Wong Kok Seng, Soongsil University, Seoul, South Korea
Mr. Praveen Ranjan Srivastava, BITS PILANI, India
Mr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong
Mr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria
Dr. Riktesh Srivastava, Skyline University, UAE
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt
and Department of Computer science, Taif University, Saudi Arabia

Mr. Tirthankar Gayen, IIT Kharagpur, India
Ms. Huei-Ru Tseng, National Chiao Tung University, Taiwan
Prof. Ning Xu, Wuhan University of Technology, China
Mr Mohammed Salem Binwahlan, Hadhramout University of Science and Technology, Yemen
& Universiti Teknologi Malaysia, Malaysia.
Dr. Aruna Ranganath, Bhoj Reddy Engineering College for Women, India
Mr. Hafeezullah Amin, Institute of Information Technology, KUST, Kohat, Pakistan
Prof. Syed S. Rizvi, University of Bridgeport, USA
Mr. Shahbaz Pervez Chattha, University of Engineering and Technology Taxila, Pakistan
Dr. Shishir Kumar, Jaypee University of Information Technology, Wakanaghat (HP), India
Mr. Shahid Mumtaz, Portugal Telecommunication, Instituto de Telecomunicações (IT) , Aveiro, Portugal
Mr. Rajesh K Shukla, Corporate Institute of Science & Technology Bhopal M P
Dr. Poonam Garg, Institute of Management Technology, India
Mr. S. Mehta, Inha University, Korea
Mr. Dilip Kumar S.M, University Visvesvaraya College of Engineering (UVCE), Bangalore University,
Bangalore
Prof. Malik Sikander Hayat Khiyal, Fatima Jinnah Women University, Rawalpindi, Pakistan
Dr. Virendra Gomase , Department of Bioinformatics, Padmashree Dr. D.Y. Patil University
Dr. Irraivan Elamvazuthi, University Technology PETRONAS, Malaysia
Mr. Saqib Saeed, University of Siegen, Germany
Mr. Pavan Kumar Gorakavi, IPMA-USA [YC]
Dr. Ahmed Nabih Zaki Rashed, Menoufia University, Egypt
Prof. Shishir K. Shandilya, Rukmani Devi Institute of Science & Technology, India
Mrs.J.Komala Lakshmi, SNR Sons College, Computer Science, India
Mr. Muhammad Sohail, KUST, Pakistan
Dr. Manjaiah D.H, Mangalore University, India
Dr. S Santhosh Baboo, D.G.Vaishnav College, Chennai, India
Prof. Dr. Mokhtar Beldjehem, Sainte-Anne University, Halifax, NS, Canada
Dr. Deepak Laxmi Narasimha, Faculty of Computer Science and Information Technology, University of
Malaya, Malaysia
Prof. Dr. Arunkumar Thangavelu, Vellore Institute Of Technology, India
Mr. M. Azath, Anna University, India
Mr. Md. Rabiul Islam, Rajshahi University of Engineering & Technology (RUET), Bangladesh
Mr. Aos Alaa Zaidan Ansaef, Multimedia University, Malaysia
Dr Suresh Jain, Professor (on leave), Institute of Engineering & Technology, Devi Ahilya University, Indore
(MP) India,
Dr. Mohammed M. Kadhum, Universiti Utara Malaysia
Mr. Hanumanthappa. J. University of Mysore, India
Mr. Syed Ishtiaque Ahmed, Bangladesh University of Engineering and Technology (BUET)
Mr Akinola Solomon Olalekan, University of Ibadan, Ibadan, Nigeria

Mr. Santosh K. Pandey, Department of Information Technology, The Institute of Chartered Accountants of India

Dr. P. Vasant, Power Control Optimization, Malaysia

Dr. Petr Ivankov, Automatika - S, Russian Federation

Dr. Utkarsh Seetha, Data Infosys Limited, India

Mrs. Priti Maheshwary, Maulana Azad National Institute of Technology, Bhopal

Dr. (Mrs) Padmavathi Ganapathi, Avinashilingam University for Women, Coimbatore

Assist. Prof. A. Neela madheswari, Anna university, India

Prof. Ganesan Ramachandra Rao, PSG College of Arts and Science, India

Mr. Kamanashis Biswas, Daffodil International University, Bangladesh

Dr. Atul Gonsai, Saurashtra University, Gujarat, India

Mr. Angkoon Phinyomark, Prince of Songkla University, Thailand

Mrs. G. Nalini Priya, Anna University, Chennai

Dr. P. Subashini, Avinashilingam University for Women, India

Assoc. Prof. Vijay Kumar Chakka, Dhirubhai Ambani IICT, Gandhinagar ,Gujarat

Mr Jitendra Agrawal, : Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal

Mr. Vishal Goyal, Department of Computer Science, Punjabi University, India

Dr. R. Baskaran, Department of Computer Science and Engineering, Anna University, Chennai

Assist. Prof, Kanwalvir Singh Dhindsa, B.B.S.B.Engg.College, Fatehgarh Sahib (Punjab), India

Dr. Jamal Ahmad Dargham, School of Engineering and Information Technology, Universiti Malaysia Sabah

Mr. Nitin Bhatia, DAV College, India

Dr. Dhavachelvan Ponnurangam, Pondicherry Central University, India

Dr. Mohd Faizal Abdollah, University of Technical Malaysia, Malaysia

Assist. Prof. Sonal Chawla, Panjab University, India

Dr. Abdul Wahid, AKG Engg. College, Ghaziabad, India

Mr. Arash Habibi Lashkari, University of Malaya (UM), Malaysia

Mr. Md. Rajibul Islam, Ibnu Sina Institute, University Technology Malaysia

Professor Dr. Sabu M. Thampi, .B.S Institute of Technology for Women, Kerala University, India

Mr. Noor Muhammed Nayeem, Université Lumière Lyon 2, 69007 Lyon, France

Dr. Himanshu Aggarwal, Department of Computer Engineering, Punjabi University, India

Prof R. Naidoo, Dept of Mathematics/Center for Advanced Computer Modelling, Durban University of Technology, Durban,South Africa

Prof. Mydhili K Nair, M S Ramaiah Institute of Technology(M.S.R.I.T), Affiliated to Visweswaraiah Technological University, Bangalore, India

M. Prabu, Adhiyamaan College of Engineering/Anna University, India

Mr. Swakkhar Shatabda, Department of Computer Science and Engineering, United International University, Bangladesh

Dr. Abdur Rashid Khan, ICIT, Gomal University, Dera Ismail Khan, Pakistan

Mr. H. Abdul Shabeer, I-Nautix Technologies,Chennai, India

Dr. M. Aramudhan, Perunthalaivar Kamarajar Institute of Engineering and Technology, India

Dr. M. P. Thapliyal, Department of Computer Science, HNB Garhwal University (Central University), India
Prof Ekta Walia Bhullar, Maharishi Markandeshwar University, Mullana (Ambala), India
Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran
Mr. Zeashan Hameed Khan, : Université de Grenoble, France
Prof. Anil K Ahlawat, Ajay Kumar Garg Engineering College, Ghaziabad, UP Technical University, Lucknow
Mr. Longe Olumide Babatope, University Of Ibadan, Nigeria
Associate Prof. Raman Maini, University College of Engineering, Punjabi University, India
Dr. Maslin Masrom, University Technology Malaysia, Malaysia
Sudipta Chattopadhyay, Jadavpur University, Kolkata, India
Dr. Dang Tuan NGUYEN, University of Information Technology, Vietnam National University - Ho Chi Minh City
Dr. Mary Lourde R., BITS-PILANI Dubai , UAE
Dr. Abdul Aziz, University of Central Punjab, Pakistan
Mr. Karan Singh, Gautam Budtha University, India
Mr. Avinash Pokhriyal, Uttar Pradesh Technical University, Lucknow, India
Associate Prof Dr Zuraini Ismail, University Technology Malaysia, Malaysia
Assistant Prof. Yasser M. Alginahi, College of Computer Science and Engineering, Taibah University, Madinah Munawwarah, KSA
Mr. Dakshina Ranjan Kisku, West Bengal University of Technology, India
Mr. Raman Kumar, Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India
Associate Prof. Samir B. Patel, Institute of Technology, Nirma University, India
Dr. M.Munir Ahamed Rabbani, B. S. Abdur Rahman University, India
Asst. Prof. Koushik Majumder, West Bengal University of Technology, India
Dr. Alex Pappachen James, Queensland Micro-nanotechnology center, Griffith University, Australia
Assistant Prof. S. Hariharan, B.S. Abdur Rahman University, India
Asst Prof. Jasmine. K. S, R.V.College of Engineering, India
Mr Naushad Ali Mamode Khan, Ministry of Education and Human Resources, Mauritius
Prof. Mahesh Goyani, G H Patel Collge of Engg. & Tech, V.V.N, Anand, Gujarat, India
Dr. Mana Mohammed, University of Tlemcen, Algeria
Prof. Jatinder Singh, Universal Institutiion of Engg. & Tech. CHD, India
Mrs. M. Anandhavalli Gauthaman, Sikkim Manipal Institute of Technology, Majitar, East Sikkim
Dr. Bin Guo, Institute Telecom SudParis, France
Mrs. Maleika Mehr Nigar Mohamed Heenaye-Mamode Khan, University of Mauritius
Prof. Pijush Biswas, RCC Institute of Information Technology, India
Mr. V. Bala Dhandayuthapani, Mekelle University, Ethiopia
Mr. Irfan Syamsuddin, State Polytechnic of Ujung Pandang, Indonesia
Mr. Kavi Kumar Khedo, University of Mauritius, Mauritius
Mr. Ravi Chandiran, Zagro Singapore Pte Ltd. Singapore
Mr. Milindkumar V. Sarode, Jawaharlal Darda Institute of Engineering and Technology, India
Dr. Shamimul Qamar, KSJ Institute of Engineering & Technology, India

Dr. C. Arun, Anna University, India
Assist. Prof. M.N.Birje, Basaveshwar Engineering College, India
Prof. Hamid Reza Naji, Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran
Assist. Prof. Debasis Giri, Department of Computer Science and Engineering, Haldia Institute of Technology
Subhabrata Barman, Haldia Institute of Technology, West Bengal
Mr. M. I. Lali, COMSATS Institute of Information Technology, Islamabad, Pakistan
Dr. Feroz Khan, Central Institute of Medicinal and Aromatic Plants, Lucknow, India
Mr. R. Nagendran, Institute of Technology, Coimbatore, Tamilnadu, India
Mr. Amnach Khawne, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, Thailand
Dr. P. Chakrabarti, Sir Padampat Singhanian University, Udaipur, India
Mr. Nafiz Imtiaz Bin Hamid, Islamic University of Technology (IUT), Bangladesh.
Shahab-A. Shamshirband, Islamic Azad University, Chalous, Iran
Prof. B. Priestly Shan, Anna Univeristy, Tamilnadu, India
Venkatramreddy Velma, Dept. of Bioinformatics, University of Mississippi Medical Center, Jackson MS USA
Akshi Kumar, Dept. of Computer Engineering, Delhi Technological University, India
Dr. Umesh Kumar Singh, Vikram University, Ujjain, India
Mr. Serguei A. Mokhov, Concordia University, Canada
Mr. Lai Khin Wee, Universiti Teknologi Malaysia, Malaysia
Dr. Awadhesh Kumar Sharma, Madan Mohan Malviya Engineering College, India
Mr. Syed R. Rizvi, Analytical Services & Materials, Inc., USA
Dr. S. Karthik, SNS College of Technology, India
Mr. Syed Qasim Bukhari, CIMET (Universidad de Granada), Spain
Mr. A.D.Potgantwar, Pune University, India
Dr. Himanshu Aggarwal, Punjabi University, India
Mr. Rajesh Ramachandran, Naipunya Institute of Management and Information Technology, India
Dr. K.L. Shunmuganathan, R.M.K Engg College, Kavaraipeitai, Chennai
Dr. Prasant Kumar Pattnaik, KIST, India.
Dr. Ch. Aswani Kumar, VIT University, India
Mr. Ijaz Ali Shoukat, King Saud University, Riyadh KSA
Mr. Arun Kumar, Sir Padam Pat Singhanian University, Udaipur, Rajasthan
Mr. Muhammad Imran Khan, Universiti Teknologi PETRONAS, Malaysia
Dr. Natarajan Meghanathan, Jackson State University, Jackson, MS, USA
Mr. Mohd Zaki Bin Mas'ud, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia
Prof. Dr. R. Geetharamani, Dept. of Computer Science and Eng., Rajalakshmi Engineering College, India
Dr. Smita Rajpal, Institute of Technology and Management, Gurgaon, India
Dr. S. Abdul Khader Jilani, University of Tabuk, Tabuk, Saudi Arabia
Mr. Syed Jamal Haider Zaidi, Bahria University, Pakistan
Dr. N. Devarajan, Government College of Technology, Coimbatore, Tamilnadu, INDIA
Mr. R. Jagadeesh Kannan, RMK Engineering College, India
Mr. Deo Prakash, Shri Mata Vaishno Devi University, India

Mr. Mohammad Abu Naser, Dept. of EEE, IUT, Gazipur, Bangladesh
Assist. Prof. Prasun Ghosal, Bengal Engineering and Science University, India
Mr. Md. Golam Kaosar, School of Engineering and Science, Victoria University, Melbourne City, Australia
Mr. R. Mahammad Shafi, Madanapalle Institute of Technology & Science, India
Dr. F.Sagayaraj Francis, Pondicherry Engineering College, India
Dr. Ajay Goel, HIET, Kaithal, India
Mr. Nayak Sunil Kashibara, Bahirji Smarak Mahavidyalaya, India
Mr. Suhas J Manangi, Microsoft India
Dr. Kalyankar N. V., Yeshwant Mahavidyalaya, Nanded, India
Dr. K.D. Verma, S.V. College of Post graduate studies & Research, India
Dr. Amjad Rehman, University Technology Malaysia, Malaysia
Mr. Rachit Garg, L K College, Jalandhar, Punjab
Mr. J. William, M.A.M college of Engineering, Trichy, Tamilnadu, India
Prof. Jue-Sam Chou, Nanhua University, College of Science and Technology, Taiwan
Dr. Thorat S.B., Institute of Technology and Management, India
Mr. Ajay Prasad, Sir Padampat Singhanian University, Udaipur, India
Dr. Kamaljit I. Lakhtaria, Atmiya Institute of Technology & Science, India
Mr. Syed Rafiul Hussain, Ahsanullah University of Science and Technology, Bangladesh
Mrs Fazeela Tunnisa, Najran University, Kingdom of Saudi Arabia
Mrs Kavita Taneja, Maharishi Markandeshwar University, Haryana, India
Mr. Maniyar Shiraz Ahmed, Najran University, Najran, KSA
Mr. Anand Kumar, AMC Engineering College, Bangalore
Dr. Rakesh Chandra Gangwar, Beant College of Engg. & Tech., Gurdaspur (Punjab) India
Dr. V V Rama Prasad, Sree Vidyanikethan Engineering College, India
Assist. Prof. Neetesh Kumar Gupta, Technocrats Institute of Technology, Bhopal (M.P.), India
Mr. Ashish Seth, Uttar Pradesh Technical University, Lucknow, UP India
Dr. V V S S Balaram, Sreenidhi Institute of Science and Technology, India
Mr Rahul Bhatia, Lingaya's Institute of Management and Technology, India
Prof. Niranjana Reddy, P, KITS, Warangal, India
Prof. Rakesh. Lingappa, Vijetha Institute of Technology, Bangalore, India
Dr. Mohammed Ali Hussain, Nimra College of Engineering & Technology, Vijayawada, A.P., India
Dr. A.Srinivasan, MNM Jain Engineering College, Rajiv Gandhi Salai, Thorapakkam, Chennai
Mr. Rakesh Kumar, M.M. University, Mullana, Ambala, India
Dr. Lena Khaled, Zarqa Private University, Aman, Jordan
Ms. Supriya Kapoor, Patni/Lingaya's Institute of Management and Tech., India
Dr. Tossapon Boongoen, Aberystwyth University, UK
Dr. Bilal Alatas, Firat University, Turkey
Assist. Prof. Jyoti Praakash Singh, Academy of Technology, India
Dr. Ritu Soni, GNG College, India

Dr . Mahendra Kumar , Sagar Institute of Research & Technology, Bhopal, India.
Dr. Binod Kumar, India
Dr. Muzhir Shaban Al-Ani, Amman Arab University Amman – Jordan
Dr. T.C. Manjunath , ATRIA Institute of Tech, India
Mr. Muhammad Zakarya, COMSATS Institute of Information Technology (CIIT), Pakistan
Assist. Prof. Harmunish Taneja, M. M. University, India
Dr. Chitra Dhawale , SICSR, Model Colony, Pune, India
Mrs Sankari Muthukaruppan, Nehru Institute of Engineering and Technology, Anna University, India
Mr. Aaqif Afzaal Abbasi, National University Of Sciences And Technology, Islamabad
Prof. Ashutosh Kumar Dubey, Trinity Institute of Technology and Research Bhopal, India
Mr. G. Appasami, Dr. Pauls Engineering College, India
Mr. M Yasin, National University of Science and Tech, karachi (NUST), Pakistan
Mr. Yaser Miaji, University Utara Malaysia, Malaysia
Mr. Shah Ahsanul Haque, International Islamic University Chittagong (IIUC), Bangladesh
Prof. (Dr) Syed Abdul Sattar, Royal Institute of Technology & Science, India
Dr. S. Sasikumar, Roever Engineering College
Assist. Prof. Monit Kapoor, Maharishi Markandeshwar University, India
Mr. Nwaocha Vivian O, National Open University of Nigeria
Dr. M. S. Vijaya, GR Govindarajulu School of Applied Computer Technology, India
Assist. Prof. Chakresh Kumar, Manav Rachna International University, India
Mr. Kunal Chadha , R&D Software Engineer, Gemalto , Singapore
Mr. Pawan Jindal, Jaypee University of Engineering and Technology, India
Mr. Mueen Uddin, Universiti Teknologi Malaysia, UTM , Malaysia
Dr. Dhuha Basheer abdullah, Mosul university, Iraq
Mr. S. Audithan, Annamalai University, India
Prof. Vijay K Chaudhari, Technocrats Institute of Technology , India
Associate Prof. Mohd Ilyas Khan, Technocrats Institute of Technology , India
Dr. Vu Thanh Nguyen, University of Information Technology, HoChiMinh City, VietNam
Assist. Prof. Anand Sharma, MITS, Lakshmangarh, Sikar, Rajasthan, India
Prof. T V Narayana Rao, HITAM Engineering college, Hyderabad
Mr. Deepak Gour, Sir Padampat Singhania University, India
Assist. Prof. Amutharaj Joyson, Kalasalingam University, India
Mr. Ali Balador, Islamic Azad University, Iran
Mr. Mohit Jain, Maharaja Surajmal Institute of Technology, India
Mr. Dilip Kumar Sharma, GLA Institute of Technology & Management, India
Dr. Debojyoti Mitra, Sir padampat Singhania University, India
Dr. Ali Dehghantanha, Asia-Pacific University College of Technology and Innovation, Malaysia
Mr. Zhao Zhang, City University of Hong Kong, China
Prof. S.P. Setty, A.U. College of Engineering, India
Prof. Patel Rakeshkumar Kantilal, Sankalchand Patel College of Engineering, India

Mr. Biswajit Bhowmik, Bengal College of Engineering & Technology, India

Mr. Manoj Gupta, Apex Institute of Engineering & Technology, India

Assist. Prof. Ajay Sharma, Raj Kumar Goel Institute Of Technology, India

Assist. Prof. Ramveer Singh, Raj Kumar Goel Institute of Technology, India

Dr. Hanan Elazhary, Electronics Research Institute, Egypt

Dr. Hosam I. Faiq, USM, Malaysia

Prof. Dipti D. Patil, MAEER's MIT College of Engg. & Tech, Pune, India

Assist. Prof. Devendra Chack, BCT Kumaon engineering College Dwarahat Almora, India

Prof. Manpreet Singh, M. M. Engg. College, M. M. University, India

Assist. Prof. M. Sadiq ali Khan, University of Karachi, Pakistan

Mr. Prasad S. Halgaonkar, MIT - College of Engineering, Pune, India

Dr. Imran Ghani, Universiti Teknologi Malaysia, Malaysia

Prof. Varun Kumar Kakar, Kumaon Engineering College, Dwarahat, India

Assist. Prof. Nisheeth Joshi, Apaji Institute, Banasthali University, Rajasthan, India

Associate Prof. Kunwar S. Vaisla, VCT Kumaon Engineering College, India

Prof Anupam Choudhary, Bhilai School Of Engg., Bhilai (C.G.), India

Mr. Divya Prakash Shrivastava, Al Jabal Al garbi University, Zawya, Libya

Associate Prof. Dr. V. Radha, Avinashilingam Deemed university for women, Coimbatore.

Dr. Kasarapu Ramani, JNT University, Anantapur, India

Dr. Anuraag Awasthi, Jayoti Vidyapeeth Womens University, India

CALL FOR PAPERS
International Journal of Computer Science and Information Security
IJCSIS 2010
ISSN: 1947-5500
<http://sites.google.com/site/ijcsis/>

International Journal Computer Science and Information Security, now at its sixth edition, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2010 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

Track A: Security

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity
Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc.), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on

its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

Track B: Computer Science

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embedded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid

Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail ijcsiseditor@gmail.com. Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at <http://sites.google.com/site/ijcsis/authors-notes> .



© IJCSIS PUBLICATION 2010
ISSN 1947 5500